DOCUMENT RESUME

ED 117 185 95 TM 005 042

AUTHOR	Young, Malcolm B.; Schuh, Russell G.
TITLE	Evaluation and Educational Decision Making: A Functional Guide to Evaluating Career Education.
INSTITUTION	Development Associates, Inc., Washington, D.C.
SPONS AGENCY	Office of Education (DHEW), Washington, D.C.
PUB DATE	Sep 75
CONTRACT	OEC-0-73-6663
NOTE	133p.
EDRS PRICE	MF-$0.76 HC-$6.97 Plus Postage
DESCRIPTORS	*Administrator Guides; *Career Education; Data Analysis; Data Collection; *Decision Making; Educational Objectives; Evaluation Criteria; *Evaluation Methods; Information Dissemination; Interprofessional Relationship; Management; Measurement Techniques; Program Design; *Program Evaluation; Program Improvement; Program Planning; Sampling; Test Reviews

ABSTRACT

This guide is intended to assist in the evaluation of career education programs. It has been developed around the concept that the evaluation should be viewed as a management tool for the improvement of program performance. The guide recognizes the key roles of the evaluators and program managers in the evaluation process. It is addressed principally to the program managers and their staffs to provide the background on evaluation techniques, issues, and problems needed for decision-making and program direction and for implementation of changes indicated by the evaluation studies. At the same time, it highlights techniques and procedures that evaluators will find useful in planning and implementing evaluation studies. The objective of the guide is to contribute to the improvement of career education efforts in local school districts. It strives to convey the importance and utility of program evaluation to the management of career education programs and to offer practical assistance in conducting such evaluations. This guide, in draft version, was published and disseminated for trial use. Hundreds of personnel provided feedback resulting in this final version. USOE undertook a review of instruments which attempted to measure career education student outcomes. For two summers a review panel reviewed instruments, and the results of these reviews are included in the appendix. (RC)

Documents acquired by ERIC include many informal unpublished materials not available from other sources. ERIC makes every effort to obtain the best copy available. Nevertheless, items of marginal reproducibility are often encountered and this affects the quality of the microfiche and hardcopy reproductions ERIC makes available via the ERIC Document Reproduction Service (EDRS). EDRS is not responsible for the quality of the original document. Reproductions supplied by EDRS are the best that can be made from the original.

ERIC

ED117185
TM 005 042

EVALUATION AND EDUCATIONAL DECISION-MAKING

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

SCOPE OF INTEREST NOTICE
The ERIC Facility has assigned this document for processing to:
TM
CE
In our judgement, this document is also of interest to the clearinghouses noted to the right. Indexing should reflect their special points of view.

A FUNCTIONAL GUIDE TO EVALUATING CAREER EDUCATION

PREPARED FOR
U.S. DEPARTMENT OF HEALTH,
EDUCATION, AND WELFARE
OFFICE OF EDUCATION

BY
DEVELOPMENT ASSOCIATES, INC.
WASHINGTON, D.C.
SEPTEMBER 1975

The work reported herein was performed pursuant to a contract with the United States Department of Health, Education, and Welfare, Office of Education. Contractors undertaking such projects under government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

ACKNOWLEDGMENTS

This guide has taken almost two years to complete and during that time numerous individuals and organizations have lent various types of support to this effort. Development Associates is indebted to these individuals and organizations for their interest and assistance.

Credit for the conception and initial support of the guide within the U.S. Office of Education belongs to Kenneth Hoyt and Gerry Elbers of the Office of Career Education, Sidney High and Joyce Cook of the Division of Vocational and Technical Education, and Alice Scates of the Office of Planning, Budgeting, and Evaluation.

The first draft was disseminated for field testing during the 1974-75 school year and literally hundreds of career education practitioners, program managers, and evaluators shared their experiences and provided inputs which were invaluable. For providing assistance during the early stages of the development of the first draft of the guide, our thanks are extended to A. Jackson Stenner of IBEX, Inc., and Jerry Walker of the Center for Vocational Education, Ohio State University.

Within Development Associates, substantive contributions were made by Stanley Lichtenstein, Robert Haupt, John Morrison, and Gerald Goldstein. Douglas Pepin provided technical guidance in the difficult areas of analysis and design. Susan Shimer and Jean Rogers provided special assistance in graphics and various other support areas. Hiram Phillips undertook the difficult task of reading and editing the text. His wide experience and sharp eye contributed significantly to the clarification and continuity of the text.

We are particularly indebted to three individuals for their support and contributions throughout the lengthy process of completing the guide.

> Terry Newell of the Office of Career Education devoted many hours to facilitating the input of career education practitioners and to reviewing drafts of several chapters.
>
> Alice Scates of the Office of Planning, Budgeting, and Evaluation provided support, insight, and professional leadership without which this effort would certainly have been weakened.

Peter B. Davis, Senior Vice President of Development Associates, provided effective leadership and contributions ranging from conceptual input to proofing text.

Malcolm Young
Russell Schuh
Development Associates, Inc.
September 1975

TABLE OF CONTENTS

		Page
Acknowledgements		
Preface		i
Chapter One:	Uses and Limitations of Evaluation	1
Chapter Two:	Plan for Evaluating Career Education	13
Chapter Three:	Preparing an Overview of the Program (Step 1)	22
Chapter Four:	Specify Participant Groups to Evaluate (Step 2)	30
Chapter Five:	Formulate Evaluation Questions (Step 3)	37
Chapter Six:	Identify Data Sources (Step 4)	42
Chapter Seven:	Select or Develop Instruments or Other Measures (Step 5)	45
Chapter Eight:	Consider Design and Analysis Alternatives (Step 6)	60
Chapter Nine:	Prepare Sampling Plan (Step 7)	69
Chapter Ten:	Finalize Design and Analysis and Prepare Reporting Plan (Steps 8 and 9)	73
Chapter Eleven:	Prepare Data Collection Plan (Step 10)	78
Chapter Twelve:	Implement the Evaluation (Steps 11 and 12)	83
Appendix:	Summary of Results: Career Education Instrument Review conducted for U.S. Office of Education	a-1
	Introduction	a-3
	Panel Procedures	a-4
	Results of the Review	a-6
	Career Education Student Objectives	a-27
	List of Panel Members	a-35

PREFACE

This guide is intended to assist in the evaluation of career education programs. It has been developed around the concept that the evaluation should be viewed as a management tool for the improvement of program performance. When this point of view is made explicit, the frequently felt resistances to evaluation are often overcome since the evaluation findings from this perspective are neither "good" nor "bad," but useful and informative.

This guide recognizes the key roles of the evaluators and program managers in the evaluation process. It is addressed principally to the program managers and their staffs to provide the background on evaluation techniques, issues, and problems needed for decision-making and program direction and for implementation of changes indicated by the evaluation studies. At the same time, it highlights techniques and procedures that evaluators will find useful in planning and implementing evaluation studies. Recognizing that program managers are not expected to be technical experts in all areas covered, a common body of knowledge for evaluators and program managers, it is hoped, will establish a base for better communication so essential to the evaluation process and improved program management.

The objective of the guide is to contribute to the improvement of career education efforts in local school districts. More specifically, it strives to convey the importance and utility of program evaluation to the management of career education programs and to offer practical assistance in conducting such evaluations. While some writing in the area of program evaluation distinguishes between process or formative and outcome or summative evaluations, the approach advocated in this guide does not fall neatly into either category. Improvement in career education requires both process and outcome information, and if these evaluation categories must be used, the approach may best be viewed as a short term "summative" evaluation conducted for "formative" purposes.

Throughout the guide an effort has been made to emphasize practical information and guidelines for use by managers of career education programs at the local level. The choice of this focus forced the consideration of several important questions; three of the most important were:

How restrictive a definition of career education should be used?

What should be the balance between advocating "sound" research procedures and setting forth common practices in educational evaluation which admittedly deviate from such procedures but generally are more possible to implement in a school environment?

How much detail or technical depth should be provided?

Obviously, these questions had to be answered. With respect to the first, we chose to define career education broadly. The career education concept has achieved prominence considerably faster than it has achieved precise definition. Since 1970, legislation supporting career education has been passed, funds have been allocated, special offices established at the state and federal levels, and hundreds of school districts have initiated career education programs. Recognizing the existence of a great diversity in program operations, we have chosen to stay at the broad conceptual level rather than base this guide on a narrow definition of career education.

Our desire to see broader use of scientific research methods has been tempered by our belief that to hold to this model as the only right way will result in lamentable delays in the spread of systematic efforts at evaluation which hold the prospect of some immediate program improvements. This same pragmatic approach has led us, in response to the third question, to provide the program manager with the kinds of information which will be useful in making decisions affecting evaluation and to recapitulate information known to many specialists in evaluation without making any effort to produce a technical manual on evaluation.

Finally, we sought to provide the program manager, presumably neither a research scientist nor a professional evaluator, with a conceptual framework, some practical advice, and a basis from which to raise relevant questions and seek specific advice. As a result, we chose to vary the level of detail, depending on what we judged to be a program manager's "need to know." For example, we provide a more extensive discussion of selecting and developing instruments than of the areas of sampling and analysis.

This guide for managers of career education programs has taken almost two years to develop. In August 1974, a draft version was published and disseminated for trial use. During the 1974-75 school year, the draft was used in many of the Vocational Exemplary Projects funded under Part D of the Vocational Education Act and in several unrelated efforts throughout the country.

The draft version was developed only after extensive input from career educators at all levels. The Office of Career Education devoted part of each of a series of 20 mini-conferences for practitioners during the summer of 1974 to evaluation. This was an important source of input from the program staff. There was also an evaluators' panel convened to provide input from those responsible for conducting evaluations, as well as a panel of nationally recognized experts in the area of tests and measurements.

After the guide's initial distribution and use, feedback was sought from as many sources as possible. The Division of Vocational-Technical Education's conference for Part D program directors provided an extremely valuable opportunity for feedback from staff, managers, and evaluators. Also, feedback was sought individually from evaluators, state education agencies, and from the U.S. Office of Education regional program personnel. By the end of August 1975, literally hundreds of program managers, evaluators, and test experts had made some input into the development of this guide.

The most noticeable change from the draft to this final version is a more explicit focus on program managers. In addition, three widely used sets of career education student outcome statements are found in the Appendix and reference to a specific set in the text has been avoided. In a comparison with the draft version, other more subtle changes in substance will be noted. These, too, resulted from the feedback of managers and evaluators who suggested areas of improvement.

Throughout, an area of special concern and attention has been the question of instrumentation. During the summer of 1974, and again during the summer of 1975, the USOE undertook a review of instruments which attempt to measure career education student outcomes. Each summer a review panel was convened for the purpose of reviewing instruments. The results of the panels' reviews are included in the Appendix.

Finally, it should be noted that the development of any system for the evaluation of specific programs is evolutionary in nature. This guide is not viewed by its developers as a final product. It is an important step in attempting to meet the management needs of career education practitioners. But experience should and will suggest improvements and refinements of this product. Not only should it be that way, but such activity is consistent with the approach to evaluation advocated in this guide.

CHAPTER ONE
USES AND LIMITATIONS OF EVALUATIONS

Introduction

Experience throughout the United States indicates that sound program management is critical to the success of career education. Because evaluation is an integral part of management, this guide is directed primarily toward those responsible for program management – managers of career education programs in local school districts and their staff – since, in practice, it is they who will make the most critical evaluation decisions. That is, they decide:

whether to have an evaluation;
who will do it;
what will be its scope and constraints; and
what will become of its results.

Although focused primarily on the concerns of the program manager, this guide will be of importance to professional evaluators and others providing specialized services to career education programs because their roles are intertwined in the process and in the results of evaluation. Much of this chapter will be given to a discussion of the roles of evaluators and program managers and their interrelationships.

Evaluation provides one of the most important bases for decisions affecting the nature, scope, and operations of the program because it tells what has happened in the past as a basis for planning and future action. Evaluation has been defined as a process ". . . designed to assist management to obtain reasonably objective information about projects and programs in a regular fashion so that lessons learned can be applied to current planning decisions and future operations."[1]

Because of this emphasis on utilization of the results of evaluation, this guide will discuss first some of the end-products of the evaluation process and some of the concerns in producing these products – i.e., the findings and conclusions. This will be done first because utilization

[1] Evaluation Handbook, Agency for International Development, Washington, DC 20523, pg. 3 MC 1026, 1 Supplement II, third printing, May 1974.

for program improvement provides the frame of reference for understanding the total process. This framework then provides the setting and perspective for later discussion of program definition, design, implementation, and reporting.

This discussion should be viewed as a guide to improved evaluation of career education, and not as a set of directives. While advocating a general approach to evaluation, some of the specific techniques and procedures should be considered as among a number of different means of achieving the goal of improved evaluation. At the same time, logically defensible local innovation and experimentation is recognized as an essential part of the evaluation process.

Dialogue Between Program Managers and Evaluators

As the quality of the evaluations and as the use of the findings and conclusions are so dependent on the proper roles of the program manager and evaluator, this subject will be developed in this section. It will be amplified later in this chapter through discussion of illustrative material affecting the utilization of evaluation. In subsequent chapters the guide will discuss the methodology and techniques utilized in the production of sound evaluations useful in the improvement of career education programs.

Because of the close interrelationships between program managers and evaluators, there must be a continuing dialogue between them. The evaluators need to know the top priorities and concerns of the program managers and incorporate them into the evaluation system. The program managers, who are not expected to be experts in evaluation, need to learn from the evaluators how the evaluation system works in theory and practice so that they can make informed judgments based on the findings of the evaluation. For effective working relationships and useful results, each must know and respect the other's role.

Several stages of the dialogue are of critical importance to the results of the projected evaluation studies and the actions to be taken on the findings. The first four points introduced below, as essential points in the dialogue, are also key aspects of the four phases of the evaluation process developed further in Chapter Two.

Program Definition

At the earliest stage, the inclusion of items related to the problems of greatest concern to the program managers will automatically increase interest in the studies and raise materially the prospect that the managers will use the findings to improve the effectiveness and efficiency of planning and management.

Design

The sharing of difficulties in developing methodology for measurement of some items may lead to improved solutions to the problems. But even when this is not possible, out of the mutual understanding developed, the program managers can be more realistic in judging the significance of the findings and taking suitable action to improve planning and management.

Implementation

With the support and understanding of the program managers, the evaluation can be conducted with greater cooperation from the teachers and other personnel and with less resistance due to concern that the study may be harmful to the program or to particular individuals.

Reporting

When evaluators know the priority concerns of program managers, they can prepare their findings so that those particular items will receive special attention. This special attention is likely to carry over to other aspects of the study. Because of their personal understanding and involvement in the system, the program managers are much more likely to use the results of evaluation; that is, to take actions leading to improvements in management and planning.

Role Definition

The common interest of program managers and evaluators in the maximum use of evaluations for the improvement of career education programs should be the point of departure for this aspect of the dialogue. The blending of their different but reinforcing responsibilities should be the main thrust of the discussion.

- **Descriptive and Evaluative**

In general, evaluators have the responsibility of program examiners, reviewers, or auditors. Like auditors of fiscal matters or bank examiners, they determine and report what and how much has happened, or has not happened, and how that relates to standards or criteria. Although evaluators and program managers work together, the findings and conclusions are the responsibility of the evaluators.

Program managers have a corresponding responsibility to define the scope of the evaluation, with inputs of course, from the evaluators; to interpret the significance of the findings and conclusions; and to take such action as they believe advisable based on the evaluator's written report within policy and administrative constraints affecting the program. While this action is discretionary on the part of program managers, they may be held accountable by their superiors and the public for inaction as well as actions based upon the evaluation.

- **Meeting of Standards**

The evaluator develops the facts on the nature and level of performance and shows how performance relates to established standards or criteria. Where measurable standards have been recognized, the evaluator reports whether or not particular activities meet the standard. In areas where standards are not precise, the evaluator can only observe, measure, and report, because he lacks an adequate yardstick for measurement. In such cases it is the responsibility of the program manager to judge the extent of conformity or deviation from the standard.

Regardless of the precision or lack of precision in the standards, it is the responsibility of the program manager to determine the significance of deviations from the standards and to determine what kind of action should be taken to improve the activity for the future.

- **The Importance of "Why"**

For program managers to decide on a course of action, they need to know what the problem is, its magnitude, what caused the problem, and why. Both the evaluators and program personnel must share in this process. The evaluators should be skilled in problem identification. They can develop or apply means of measurement to many problems. By checking with program personnel during the course of the evaluation, they can often obtain adequate explanations of what caused the problem. For example, when tests show a low rate of learning, they can determine that a particular topic was not taught or was touched only briefly in the classroom. Or they might find that the explanation for a decrease in industry visits was due to a strike of school bus drivers. For many types of corrective action, these explanations are sufficient for the program manager.

In other cases, the "why" may be much more complex and yet, knowing the "why of it" may be essential to choosing among alternative courses of action. In such cases the program manager and staff may be more effective in developing the information because of their close and continuing relationship with project personnel. They might find, for example, that the reason a particular topic was not taught was because teachers did not feel competent in the use of curriculum guides, or the subject was crowded out by competing topics, or that the principal felt that that particular topic was not important and should be dropped. Obviously, the type of action required would be significantly different depending on the "why" of the action; program personnel, out of their close relationship with the project, may find it easier than the evaluator to develop this sometimes sensitive information.

- **Grant-Condition Evaluations**

In some instances evaluations may be required as a condition of an outside grant of funds. Although such evaluations may have their own objectives, they should address matters of concern to the program managers. The earlier principles of dialogue between evaluators and program managers hold here, too. To the extent that there can be a dialogue between program managers and evaluators on problem definition, design, implementation, and reporting, the utility of the evaluation is likely to be enhanced for grantor and grantee alike.

In the following section we draw on lessons from experience to illustrate further and reinforce some of these principles and thus provide further guidance for program managers and evaluators.

Some Uses of Evaluation

An outside evaluator concluded a recent report by stating that the general approach taken by the Career Education Project was good and recommending that the program be expanded within the school district. Such judgmental conclusions may not be useful to program managers, and could be harmful to the role relationship discussed above.

For a third party to recommend expansion of one program without study of the broader policy and management issues involved is to ignore the very difficult consideration of public policy and allocation of limited resources among competing program activities. Such considerations are obviously the province of policy making bodies and their managers. Had they felt the need of advice in this area they would have specifically asked for it from someone they considered competent to give it after undertaking a study of the issues involved. Similarly, for the evaluator to make the pronouncement that the general approach is "good" (or "bad") is to go beyond the scope of the evaluation unless the evaluator had the opportunity for a study in greater depth than is normally expected. Goodness is generally based upon considerations beyond ability to meet stated objectives. For example, inefficiently meeting objectives would not be as good as efficiently meeting them; the degree of goodness may also be influenced by the numbers of students affected by one program as against another.

In summary, the evaluation study is enhanced and the probabilities of its use increased when conclusions are based on and supported by the findings. Judgmental conclusions beyond the scope of the study may divert attention away from well-documented findings and conclusions, possibly weakening the impact of the evaluation.

On the other hand, program managers are likely to find evaluation reports useful when findings show that the program:

met all of its objectives or performed as expected;
met none of its objectives or failed to perform as expected; or
met some of its objectives or performed only partially as expected.

Each of these general findings offers valuable insight for program managers interested in improving program performance. Some suggestions on use of findings in each of these categories are offered below.

The program was found to have met all of its objectives or to have performed as expected.

Evaluation reports are seldom read from the first page to the last in sequence. This is especially true when managers read reports on their own projects. The natural inclination is to read the conclusions first in an attempt to find out how well the evaluator understood the program, or to determine what lies ahead in explaining the findings to the superintendent of schools.

It is not difficult to understand what managers feel when the report states that the program met all of its objectives. Relief! For a multitude of reasons, program evaluations tend to produce anxiety, and when a program has been found successful in meeting its objectives, management is spared many of the burdens that befall managers of less successful programs. But, after the anxiety has abated, the question remains of using the results for further improvement of the program.

The findings suggest, in this instance, that the program's resources were both adequate and appropriate to produce the planned results. In other words, after a period of operation it has been demonstrated that the program activities will result in the planned program outcomes. Further, it has been demonstrated that the level of program activities undertaken is adequate to produce the planned results.

With this understanding based upon past performance, there remains one important unanswered question: was the approach used the best one or is there a better way of accomplishing the same thing? In seeking to answer this question, managers are dealing with questions of program efficiency. For example, the program would be more efficient if it could serve a greater number of students and produce the desired changes in them without requiring additional resources. Program efficiency would also be improved by serving the same number of students with the same results but with a lower level of resources.

In searching for more efficient means of reaching program objectives, it is sometimes helpful to list those factors deemed to be main contributors to the program's success, ranking them from most important to least important. Beginning with the lowest ranking items, the list should be reviewed to determine which ones made little or no contribution to the success of the program and what would happen if they were eliminated. The identification and elimination of unnecessary activities is one way of improving the efficiency of a program. As stressed throughout this guide, this technique, like many others cited, is most effective when there is broad participation in the process and a consensus is developed on the desirability of improving the program by eliminating non-essential activities.

Beyond the elimination of low priority activities, there is, of course, the further need to examine the efficiency and effectiveness of ongoing activities and the points of interaction between and among different activities.

The program was found to have met none of its objectives or to have failed to perform as expected.

When an evaluation report indicates that a program failed to meet any of its objectives, it is simply describing past events. The evaluation process, however, is not complete until the reasons for this apparently negative picture are understood. In other words, the evaluation process must describe not only what happened but also why it happened (or did not happen). In seeking to explain program performance which did not meet expectations, there are some questions which often provide insights in developing explanations and planning corrective actions.

Were the project objectives obtainable in the first place?

Frequently, project objectives are laudable in their intent, yet impossible to reach in practice. One of the most vivid examples involves a program which stated that it was going to place every graduating senior in either employment or post-secondary education. No one would argue about the nobility of the cause, but in practice, the program encountered some difficulty, especially as the area in which the project was located was economically depressed (indeed, unemployment approached 25% because of an industrial shutdown). The project, of course, failed to reach its stated but unrealistic objective even though it made a substantial number of placements. Had the objective been stated more realistically, the placement effort might have been considered successful, given the condition of the labor market. In such situations one possible course of action for the program manager is to reexamine and restate this particular program objective.

Were the program's resources and inputs sufficient to produce the expected results?

It is not unusual to find a career education program failing to reach its objectives for no other reason than that too much was expected of limited resources. One is reminded of the bewildered and even heart-broken expressions on the faces of a program staff who, after a highly successful two-year career education demonstration effort in two elementary schools, attempted to expand their activities to 26 additional elementary schools. At the end of the third year, the evaluators reported that they were unable to detect any significant changes in the new schools. It is not difficult to understand what happened. In their enthusiasm to replicate the success at two schools, the program staff had attempted a dramatic increase in the range of the program without a corresponding increase in program resources. They were not, in reality, replicating the successful effort in other schools, but instead, trying to install a very limited version of the successful effort and expecting similar results.

With this kind of information available, the program manager can scale down the objectives, provide for their achievement over a longer time period, expand the resources for this activity, or develop some combination of these approaches.

Were the program activities appropriate for reaching the intended objectives?

Programs are generally evaluated to determine how well they did what they were supposed to do. It is not unusual to encounter situations in which a program undertook a set of activities, was effective and efficient in carrying them out, and at the same time, made no progress toward realizing the stated objectives. In other words, they did not do what was expected even though what was done was done well.

This is illustrated by the example of a project which was expected to increase students' career awareness and knowledge about jobs through the infusion of career education concepts and materials into the regular classroom curricula. The project used its resources for the development of high quality curriculum guides which integrated career education principles and information into the subject matter very effectively. The evaluation showed no change in students' awareness and knowledge. Why? Analysis showed that the guides had not been used by the teachers who were without training or instruction in how to use them. The project had erroneously assumed that once the guides became available they would be used immediately by the classroom teachers. Project staff may have worked hard but as yet had made no impact on students. The evaluation, however, was not concerned so much with good work as it was with demonstrable progress toward objectives.

Again, as a result of the findings of the evaluation and with careful inquiry into the "why," a program manager was able to improve the effectiveness of the project by making sure the teachers made proper use of the guides.

The program was found to have met some of its objectives or performed only partially as expected.

At a recent conference, a project director was overheard to quip that "evaluators were like teachers who never gave A's." If they couldn't find something to criticize, "they hadn't done their job." While the truth of the statement may be open to debate, it is based upon the observation that evaluators frequently conclude that programs have been only partially successful in realizing their objectives. When evaluators report such findings, managers are faced with a range of considerations in determining how best to improve their programs. It is clear that some things worked and some did not.

Typical of the "mixed bag" is a recent study in which the evaluator discovered that the teachers on one project school had participated in many more inservice training sessions than the teachers in another school. He was able to document the dramatic differences in performance of the two groups. Inquiry showed that the principal of the first school encouraged participation in the training and was a strong advocate of career education.

In this type of situation program managers try to build on successful experience. Knowing what happened and some of the reasons for it, they can plan a strategy to spread this experience to the lagging schools. Or, at the other extreme, they might conclude that the chances for success in some schools are so slight that particular activities should be dropped until circumstances change.

In planning courses of action where the evaluator reports mixed findings, the program manager may find it helpful to go through a six-step analytical process:

> Step One - - Separate the project activities into two categories: those identified as meeting objectives or performing as expected and those failing to meet objectives or failing to perform up to expectations.
>
> Step Two - - Examine the activities meeting their objectives to determine whether they can be improved, extended to reach more students, or operated with greater economy of resources as discussed earlier.
>
> Step Three - - Analyze the evaluator's findings with respect to the activities that have failed to meet objectives against the criteria discussed previously to determine what actions might be taken to overcome the reported shortcomings.
>
> Step Four - - Compare activities and categories of activities in an effort to identify the key differences which affect success and failure.

Step Five -- Plan strategy and actions to improve those activities which seem to have a reasonable chance of success, including timetables and bench-marks for periodic review of progress being made.

Step Six -- For those activities which seem to have little chance for success, consider and decide among a variety of alternatives:

- substantial modification of the activity;
- continued operation with a lower level of resources;
- shifting the resources and activities to other areas; or
- dropping the activities and using the resources elsewhere.

Conclusion

The discussion and suggestions contained in this chapter have not been offered as complete or all-inclusive. Rather, it is hoped that they will encourage program managers to view evaluation as a tool to be used in improving the quality of career education programs. When evaluation results indicate that programs have performed as expected, management should use the results to help seek ways of further improving productivity. When the findings reveal that programs did not perform as expected, they can be used to determine what is necessary to improve performance. Viewed in this manner, evaluation results should not be considered as threatening to the project but as an essential part of the process of improving the quality of career education efforts.

In the remaining chapters of this guide, a thirteen-step process for conducting evaluations is presented. The chapter which follows presents an overview of the process. Subsequent chapters discuss in more detail each of the steps.

CHAPTER TWO

PLAN FOR EVALUATING CAREER EDUCATION

Introduction

Evaluation is not an end in itself. It can serve many purposes - - the appraisal and valuation of a diamond, the judgment of the significance of results of a scientific experiment, the determination of the value of a used car, the appraisal of the effectiveness and efficiency of a program, among many others. In all of these many uses there are two common elements:

> evaluation is conducted for purposes of decision-making and, therefore, must be responsive to the needs of the decision makers; and
>
> evaluation must be made against some framework, standards, or criteria.

In this guide, evaluation is considered as part of the management of educational systems, and more particularly, part of the process of management of career education with the underlying purpose of improving program performance.

Because evaluation is part of the process of managing career education, the evaluation process must be based on the elements considered important by the decision makers, i.e., the program managers. Generally, these elements will be those that help the managers judge the effectiveness, efficiency, and costs of the program. Usually included are such matters as the numbers and types of students participating, the types of activities in which they participate, the levels and frequency of participation, what the students learn, how much they retain and apply, how much these activities cost, and many more. The numbers associated with these elements, however, are not significant in themselves. Four hundred ninth grade students visited three plants. Fourteen sixth grade classes had visitors who discussed their occupations. Are these figures good or bad? There is no way for a program manager to know what they mean unless he has some framework for judgment that permits him to say they are good or bad in comparison with something.

Framework for Evaluation: The Program Plan

The most useful framework for evaluation is a program plan that sets goals, objectives, and targets, lays out the resources that will be needed, outlines the courses of action required, sets the time frame, and describes the results expected in career education. With such a framework, the program manager is in a position to say whether some numbers have any significance for evaluation, whether the numbers show good progress, poor progress, or that the program is standing still.

A program plan is essentially a projection of the future. It starts with where we are now, what we did in the past, what we plan to do in the period ahead. It is not just a wish or a hope, but a statement of what we expect to accomplish and what we will need to do it.

Experience tells us that at times, especially the first time or two a program plan is developed, performance will fall far short of the mark. Then the planning and evaluation process - - part of good program management - - calls for a review of the program plan, the implementation process, and the evaluation system. The basic questions in this review are: What worked? What went wrong? Were there problems of implementation that had not been anticipated? Were the available resources less than expected? Were the problems more difficult than anticipated? Were the goals unrealistic? Were the tools of measurement inappropriate or faulty?

By working back and forth among program planning, implementation, and evaluation (Figure 1), the plan will become more accurate, more realistic. In the planning process itself, more attention will be given to matters of policy that must be resolved in setting goals; actions that must be taken to gain necessary resources for the program; the time needed for getting the cooperation of busy teachers; and the complexity of developing reliable indicators of career education learning, retention, and use. It is out of this process that a more useful framework for evaluation and a program of implementation that is increasingly responsive to local needs and resources are developed.

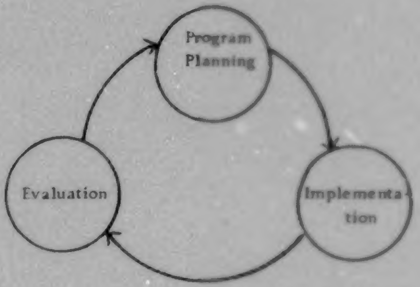


FIGURE 1. Program Management Process

Evaluation as viewed here is considered to be a continuous or cyclic process. The results of an evaluation are expected to provide the basis for program modification requiring further evaluation.

Planning the Evaluation

At the broadest level, there are four phases in evaluating any program: specifying what is to be evaluated; designing the evaluation plan; implementing the plan; and reporting the findings (Figure 2).

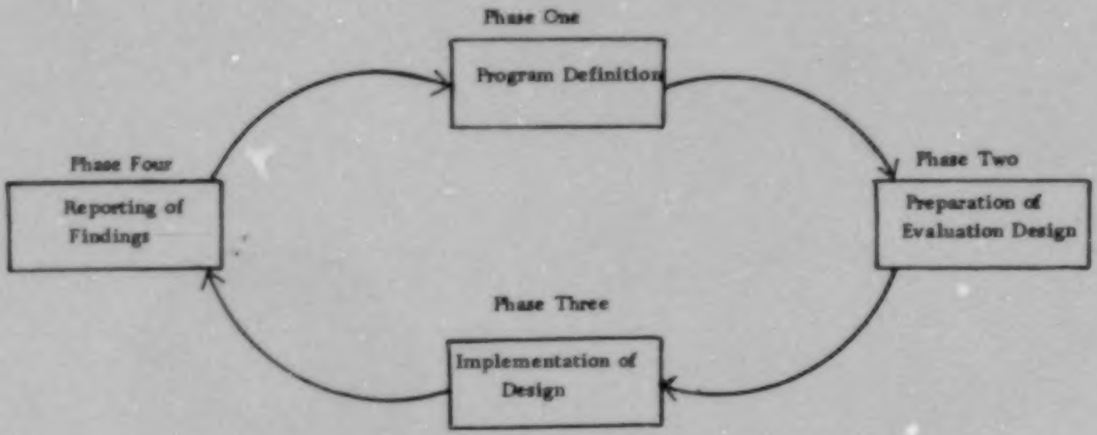

Phase One
Program Definition
Phase Two
Preparation of Evaluation Design
Phase Three
Implementation of Design
Phase Four
Reporting of Findings

FIGURE 2. Phases in an Evaluation

Within each of these four broad phases of program evaluation several steps should be completed in sequential order. More specifically, the steps which make up the program definition and design phases define (and therefore must precede) the remainder of the effort. Thus, these initial phases must make clear what is to be evaluated, why it is to be evaluated, how it is to be evaluated, and how the findings are to be reported. Because they are so crucial to the overall process, they are the focus of the major portion of this handbook.

While the entire process is important, the clarification of basic implementation and evaluation issues at the start of the process cannot be overstressed, particularly because the term career education is subject to multiple definitions. This need to specify the scope of the program has been a recurring problem in both the implementation and evaluation of career education efforts since 1970.

Many of the objectives of career education may overlap with objectives of other programs or emphases of a school district, and activities designed to lead to achievement of career education objectives may be supported financially from several sources and staffed by personnel from various areas in the district. In some situations it may be appropriate to define the program by particular funding source or combination of funding sources. For example, if a school district is supporting career education through funds received from a federal grant, state career education funds, and local revenues, it may be appropriate to include all activities regardless of funding source. What is most important in terms of the evaluation design is that this decision be made explicit at the start of the design process.

Of equal importance to the design is making explicit those broad program activities which will be included in the evaluation. For example, most school systems have taught vocational skills courses for decades. In many districts work experience (cooperative education, distributive education, etc.) courses have been ongoing for years. Clearly, vocational skills courses and work experience programs address many of the objectives of career education programs and thus legitimately might be included in the scope of the evaluation. In some districts the quantity of effort has remained constant for many years but substantive changes have occurred recently and may be associated with career education. In other districts there have not been changes either in the quantity or quality of vocational courses or work experience programs but since they are conceptually related to a comprehensive career education program they are defined by the district as part of their career education efforts. In designing the evaluation it is essential that it be specified at the outset whether such activities are to be included in total, included in part, or excluded from consideration.

Closely related to the need to specify which, if any, long-standing efforts of the school district will be included in the scope of the evaluation, is the need to specify what new activities are to be included. If the primary focus of the evaluation is in terms of funding sources or use of specific school personnel, then a decision must be made regarding the inclusion of related activities. For example, if the evaluation is to focus on a particular program supported by a specific grant, should additional activities also be included in the evaluation that were by-products of the particular program but are not supported by the grant?

Steps in Evaluation Process

To provide an overview of the entire process of evaluating career education in the context of a local school system, Figure 3 specifies 13 steps or tasks which should be completed. The first two of these steps may be considered as the pre-design or program definition phase (Phase One, page 18). Steps 3 through 9 are those necessary to complete the evaluation design (Phase Two, page 18), as they specify a procedure to solve the evaluation problem. Steps 10 and 11 relate to implementation of the design (Phase Three, page 19). Steps 12 and 13 (Phase Four, page 19) involve reporting the findings on which program modifications may be based and the use of those findings, leading to renewal of the evaluation process.

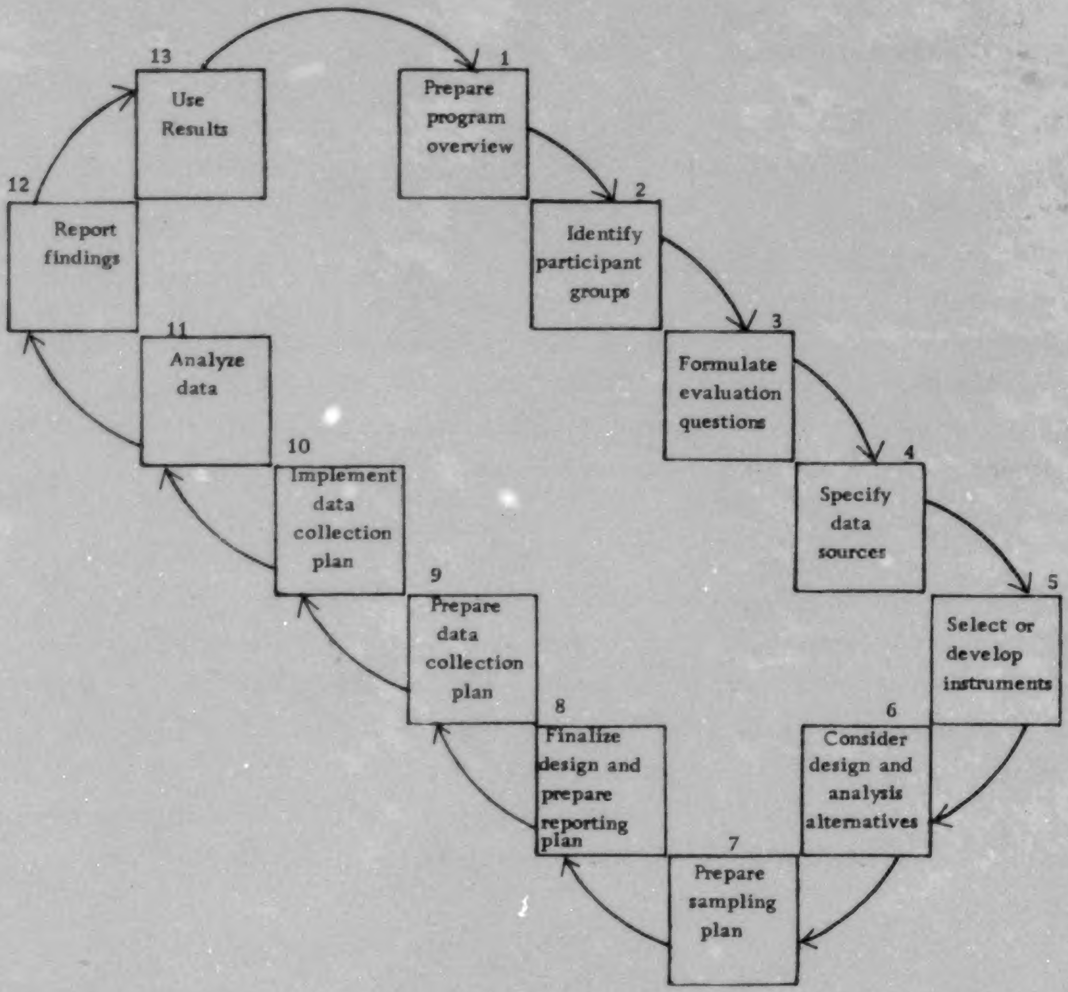

1 Prepare program overview
2 Identify participant groups
3 Formulate evaluation questions
4 Specify data sources
5 Select or develop instruments
6 Consider design and analysis alternatives
7 Prepare sampling plan
8 Finalize design and prepare reporting plan
9 Prepare data collection plan
10 Implement data collection plan
11 Analyze data
12 Report findings
13 Use Results

FIGURE 3. Overview of Steps in Evaluation Process

A narrative overview of the 13-step process is provided below:

Phase One - Program Definition

The first step is to prepare an overview of the program. The overview is a relatively brief description of the career education program being evaluated. It is useful not only for the evaluation but also in program planning and for public information purposes. Preparation of the overview is the subject of Chapter Three.

The second step is to identify discrete groups of program participants. Completion of this step requires relating program objectives, activities, and individual participants. Essentially, this is a further specification of the narrative overview. The process of identifying participant groups is the subject of Chapter Four.

Phase Two - Evaluation Design

The third step is to formulate evaluation questions. Questions should address the areas of program objectives, program activities, and program management. Both descriptive and explanatory types of questions should be asked. The process of formulating evaluation questions is the subject of Chapter Five.

The fourth step is to identify the data sources to be used in answering the evaluation questions. Possible sources of data are: those affected by the career education program, those who cause the effects, and observers of the effects. Data may be gathered directly from individuals or from school files. The identification of data sources is the subject of Chapter Six.

The fifth step is to select or develop instruments. Types of instruments include: standardized tests or questionnaires; specially developed tests or questionnaires; observation scales; interview schedules; document review summary forms; and staff reporting forms. To answer several student outcome questions, a national search for standardized instruments was conducted in the course of preparing this guide. The instruments were then systematically reviewed by panels of test and measurement experts. The results of that review and specific information about the instruments judged by the panels to be potentially useful are reported in the Appendix. In Chapter Seven the step of selecting or developing instrumentation is discussed in more generic terms.

The sixth step is to consider various evaluation design and analysis alternatives. The issues of design and analysis are thoroughly intertwined with those of data sources, instrumentation and sampling. While tentative decisions should be made at this point, final decisions must await consideration of sampling issues and a review of issues relating to sources of information and instruments. Several designs which may be considered are discussed in Chapter Eight.

The seventh step is to prepare a sampling plan. The details of it depend on decisions made in preceding steps as well as administrative considerations. Several points to consider and some general "rules of thumb" are presented in Chapter Nine.

The eighth step is to complete the design and prepare a reporting plan. All decisions made to this point are reviewed for technical feasibility, cost, and the needs of users of the evaluation findings. Once this is done and modifications are made as necessary, a plan is prepared for reporting the findings to the various users, along with target dates. This process is the subject of Chapter Ten.

The ninth step is to prepare a data collection plan. Here the tasks, personnel, and target dates for implementing the design are specified. Preparation of a data collection plan is the subject of Chapter Eleven.

Phase Three - Implementation

The tenth step is to implement the data collection plan. It is presumed that problems will arise and all will not go as smoothly as envisioned. It is important to record all deviations from the plan as an aid in analysis and interpretation.

The eleventh step is to analyze the data collected. The first task is processing and tabulating the data. Then suitable analytical techniques are selected and analysis is performed. Steps ten and eleven, as well as reporting, are the subject of Chapter Twelve.

Phase Four - Report Findings

The twelfth step is to report the findings and make recommendations as appropriate. The report should be prepared in sections on the assumption that not everyone will have need for all sections of the report. As an aid in comparing the findings of evaluations across projects, the report should provide basic data indicating the context in which the career education effort was carried out.

The final step is to use the findings of the evaluation. This has been the purpose of the entire process. Because of its importance, it was discussed in Chapter One.

Design Worksheets: An Aid to Planning

As an aid in implementing several steps in the evaluation plan, Table 2.1 contains an evaluation design worksheet. This worksheet is designed to provide a summary of the evaluation plan in a convenient and logical format. A final task in the evaluation design phase should be to complete this worksheet or some similar aid. The worksheet is provided here at the beginning of the discussion to give the reader a preview of the system.

It must also be emphasized again that the worksheet, as well as the other material in this guide, is intended to provide the manager with the background for decision-making. The program manager has no need to master all the technical matters involved in evaluation; that is the province of the evaluator. However, the program manager does need enough understanding of techniques, methodology, and problem areas in evaluation to use the findings wisely for sound decisions to improve career education.

TABLE 2. 1

EVALUATION DESIGN WORKSHEET

COLUMN 1. Program Objectives (See Chapter Four)	COLUMN 2. Related Evaluation Questions (See Chapter Five)	COLUMN 3. Specify Data Source(s) (See Chapter Six)	COLUMN 4. Instrumentation (See Chapter Seven)	COLUMN 5. Sampling (See Chapter Nine)	COLUMN 6. Design and Analysis (See Chapters Eight and Ten)
List program objective #1.	Outcome: State outcome question	List sources from whom data is to be collected	Cite Instrument or describe procedure	Specify sampling procedure and size	Indicate type of design and Specify descriptive and inferential statistical procedures
	Activity: State related activity questions	List sources from whom data is to be collected	Describe type of instrument or procedure	Specify sampling procedure and size	
	Relational: State question relating outcome to activities	Not Applicable	Not Applicable	Not Applicable	Specify Analysis Techniques
List program objective #2.	Outcome:				
	Activity:				
	Relational:				

CHAPTER THREE

STEP 1: PREPARING AN OVERVIEW OF THE PROGRAM

Introduction

An overview is a brief statement of the purposes of the career education effort, the major activities to be undertaken, the resources to be used, and the results expected. While useful f[illegible] many purposes, it should have a particular focus on planning, implementation, and evaluation, the three main elements of program management.

Preparation of the overview, if one is not available, thus becomes the first step in the evaluation of a career education program. It is an expression of the goals set for the program, the inputs required, and the outputs or results expected within a particular, specified time period in all or part of the school system. Evaluation then becomes a matter of attempting to measure and judge how much of what was planned has actually happened. When these results are known, judgments can be made of actions required to improve the program.

The Overview

When a current program plan is available, the development of a concise, written overview, usually no more than six to eight pages double-spaced, is a relatively simple task. It gives the manager a comprehensive picture or outline of the program and thus provides the context for subsequent steps in planning, implementation, and evaluation. The overview, in essence, is a summary of the program plan.

Essentially, the overview serves to clarify the manager's own view of what is to be managed and then serves as an outline for use in the evaluation. The manager may, in addition, find the overview useful for other purposes: as a summary of the program for program staff, school officials, and the general public; and, as a part of the final report. The purpose of the overview, then, is to describe in simple language what the program is and what it is not.

In serving these various ends, the overview should address, in broad terms: 1) the purposes of the career education effort; 2) the major activities to be undertaken to achieve these purposes; 3) the resources to be used; and 4) the results expected.

It may be helpful to conceive of the overview as an executive summary to be read by the school superintendent, the general public, and interested officials at the national level. In summary form, it should include the following information:

> *Scope of program* - geographic area, proportion of schools and students, curriculum areas, and grade levels covered by the program;
>
> *Origins of the program* - background of the entire program, its original and current resources and purpose;
>
> *Program goals and objectives* - present goals and objectives and their interpretation by the staff; program priorities;
>
> *Staff and program participants* - number and types of staff members in the program with summary of principal responsibilities; number of students and other participants by major subgroups such as school and grade level;
>
> *Major program activities* - current and projected activities showing staff and participant involvement; the relationship between activities and programs, and their objectives;
>
> *Funding and expenses* - identification of source(s) of funding and amounts; major non-cash contributions; major types and amounts of expenditures to provide an understanding of the size of the program and its emphases.

To repeat: the overview should not be lengthy, about six to eight pages double-spaced. It should address major points. Much of the detail can be developed elsewhere, notably in Step 2 which is discussed in Chapter Five. It should give the reader a quick picture - - an overview of what the program is, the area of its application, its costs, and accomplishments. No hard and fast rules can be given for the length of the document, but a few rule-of-thumb pointers may be helpful. Scope of the program should be a matter of one or two paragraphs. The origin of the program usually can be treated adequately in a paragraph or two. Program goals and objectives would be a longer section, about a page or two double-spaced, as only the four or five highest priority items should be included and each one summarized in a few lines. Staff and program participants can be adequately covered in no more than two paragraphs as the detail can and should be developed in Step 2. Funding and expenses also can be treated in no more than two paragraphs. Major program activities would parallel goals and objectives in length, about a page or two double-spaced, with a few sentences or a paragraph sufficient for each of the four or five major activities.

Developing the Overview

General Considerations

The timing and method of preparation of the overview merits particular attention. When prepared at the beginning of the program it should be based on expected inputs, processes, program objectives, and projected results. When updated annually, the overview should show actual inputs and results for the year completed compared with the projections previously made for that same period and forecasts or projections for the year ahead. In this way the overview becomes a tool for planning as well as for evaluation.

Original research should not be necessary in the preparation of the initial overview or the updating. Nevertheless, it is obvious that a flow of information from a variety of sources normally used by a program manager will be necessary to make it useful to him and others within and outside the school system. Grant proposals and annual reports, for example, provide excellent source material, but of course they are intended for other purposes and should not be used instead of the especially prepared overview.

Preparation of a Program Plan

When there is no reasonably current program plan, the preparation of the overview obviously will involve more work than summarizing and highlighting an existing plan. Because a program plan is useful not only in the evaluation process but serves as one of the major instruments for program management, the manager may feel it an economy of effort to develop a program plan and an overview concurrently.

As mentioned previously, the program plan provides the framework for the evaluation and the overview is essentially a summary of the plan. We cannot tell whether we have arrived or how far we have gone, or whether we are ahead of or behind schedule, or whether our costs are reasonable, unless we have some plan with various goals and realistic ways of reaching those goals. Unless, for example, a plan called for 50 percent of the students in the tenth grade in a certain area to know the job entry requirements in selected major industries, it would be hard to say whether demonstration of this knowledge by 20 percent, 40 percent, 50 percent, or 80 percent of the students represented reasonable performance. To cite a further example, an evaluation might show that 50 teachers participated in inservice training. In the absence of a plan which said that to achieve a particular goal some number should participate in certain kinds of training, it would be hard to know whether the 50 represented poor, satisfactory, or outstanding performance.

The literature relative to program planning suggests that planning, in general, involves the following elements:

- identification of a problem or need;
- assessment of the magnitude of the problem (needs assessment);
- inventory of the resources available with which to overcome the problem or meet the need;
- development of specific program objectives;
- identification of the target population;
- specification of program activities; and
- evaluation of the results.

As a practical matter, program managers seldom have the opportunity for planning in the ideal sense. The program manager is more likely to become involved in planning and evaluation of career education at some midpoint in the cycle rather than at the beginning. The school district, for example, may have already made a commitment to implement a career education program, objectives may have been set with or without formal identification of needs or an assessment of the magnitude of the problem, budgetary decisions may have already been made. In such circumstances, the manager may be faced with the design and implementation of a new project, or with managing a program that has already begun. In either event, the problem for the manager is basically the same: how to implement the project so that it will be a success.

Program experience suggests that early program planning, from a management perspective, is an important element of success. Even when prior decisions seem to limit alternatives, the successful manager pays close attention to the key elements in the planning cycle and takes whatever action may be appropriate to increase the chances of program success. To illustrate, if the review of available resources shows they are not adequate, it may be possible to persuade policy makers either to provide more resources or to establish more modest goals. Further, a careful assessment of ongoing activities may show that some needs are already being met in other ways and that modifications should, therefore, be made in the program. Thus, to the extent possible, the successful manager strives to make program resources and expectations compatible.

Of course, early planning is no guarantee of positive results, but it will substantially improve the prospects of positive performance and evaluation findings.

Key Considerations

The availability of a program plan, although valuable in itself, of course, is not essential to the preparation of an overview. The overview can be developed through review and summarization of various program documents and reports and decision-making on program direction, goals, and resources.

Prior to actually writing the overview, two preliminary sub-steps should be performed: (a) review existing statements of program objectives; and (b) review expected activities and resources. This review will facilitate writing the overview and several of the other steps in the evaluation process.

a. Review Program Objectives

Program objectives are frequently stated as broad, general goals which do not indicate specifically what the program is to accomplish or how accomplishment will be measured. For example, it is not unusual to find objectives such as: "the program will expose elementary students to career education concepts through infusion into the regular curriculum," or "students will increase their self-awareness."

While these may be desirable goals from a policy point of view, they are not very helpful to managers attempting to implement and evaluate programs. They do not answer certain necessary questions which must be answered if managers are to steer the program on a success-producing course.

Neither of the examples specify exactly which students will be affected. The first statement does not explain which grade levels will be affected and in which schools. In the second example, it is not known just what is meant by self-awareness, and therefore, what the student is expected to gain remains unstated and unclear.

On the other hand, program objectives should not be too specific since this tends to create a "laundry list" of ends which is of little or no help to anyone; persons outside the program are overwhelmed and program staff is unnecessarily constrained. In the overview, a middle ground should be sought.

Out of this early and broad based formulation of operational objectives can come the type of statements of objectives for which program managers can reasonably be held accountable and for which evaluators may be able to develop methods of measurement. Thus, the reference to self-awareness might be restated to read: "Elementary students participating in schools A, B, and C will demonstrate improved understanding of their own interests, aptitudes, and abilities." Such a statement defines self-awareness in terms that lend themselves to multiple measures (interests, aptitudes, and abilities) in the final evaluation.

Experienced evaluators expect frequently to encounter generalized statements of program objectives when evaluating educational programs. They understand that they cannot answer how well a program met a given objective when it is so stated and are trained to restate program objectives in terms amenable to measurement or verification. While outside assistance may be helpful in formulating objectives, the wise program manager will not turn the problem entirely over to someone else. When this happens, the objectives of the program are really being set by persons with neither legal nor operational responsibility for their achievement.

b. Review Program Activities and Resources

The task of planning for implementation and evaluation is not completed with the review and restatement (if necessary) of program objectives. Managers must next analyze, and in some cases specify, the program activities that relate to each objective and the resources available to carry them out. At this point, it is necessary to determine if the activities and resources are appropriate and sufficient to produce the expected outcomes. If they appear appropriate, implementation may begin with the confidence that conditions for success have been maximized. If they do not appear appropriate, managers must either admit failure from the start or modify the objective or the activity. Assuming the first option is unacceptable, modification usually requires negotiation with school administration, policy bodies, or funding sources, or some combination of any or all. The result of such negotiation should be a workable implementation plan and a reasonable concept in the overview.

Often, managers will find that the activities appear appropriate to reach the stated objectives but the resources allocated appear to be insufficient. The frequency with which this occurs is explained in part by the general nature of the decision making and funding processes. Funding sources, quite naturally, seek the broadest possible program for the smallest dollar investment. Policy boards and administrators seeking needed resources state or imply the broadest plausible program in an attempt to qualify for the desired resources. Frequently, the scope of the program is stated in general and not specific terms such as "elementary students will be involved . . ." Such statements imply total elementary level involvement but do not actually specify it.

In any event, managers faced with limited resources must review the stated program objectives to determine whether thay can be reached with the allocated resources. When it is judged not reasonable to expect to reach the stated objectives with the resources allocated, there are several techniques which managers can use simply or in combination to improve the chances for a successful program.

Restate the objectives so that they are reachable with the resources allocated. This may sometimes be accomplished by simply rewording objectives from generalizations to specifics. For example, continuing with the self-awareness objective, the restatement may indicate that "sixth grade students participating in the program will be able to identify and describe their own interests, aptitudes, and abilities." Using this technique, different kinds of limits can be established. The grade levels affected may be specified or the specific schools identified (and limited) or even a combination of grade level and school limits specified.

Develop a set of priorities. Using this technique, managers may emphasize the reaching of one or more objectives at the expense of others due to resource limitations. In employing this technique managers allocate sufficient resources to reach the highest priority objectives and devote the remaining resources toward lesser priority objectives.

Time-phase project activities. This technique requires careful planning and generally a project with a rather long term. In this instance managers devote all of a project's resources toward the realization of a limited number of objectives within a specified and limited period of time. At the end of that period the resources are redirected toward reaching another set of objectives, etc.

These techniques lead to policy and program commitments and establish areas of responsibility for which the program manager can be held accountable and, therefore, should be reflected in the overview. Consequently, the importance of putting any decisions such as these in writing cannot be overstressed. This is necessary in order to insure that policy makers, staff, evaluators, and other interested parties understand the parameters of the project.

Summary

We have, in this chapter, discussed the major purposes of the overview and important steps in its preparation. We have stressed the importance of its brevity and its clarity. In the search for clarity we have touched on a number of areas of important detail which will be treated in greater depth in later chapters. We have introduced them here to show that in producing the statement of broad general scope that is the overview, it must be based on solid analysis and consideration of the overview as a part of the whole evaluation process.

Once the overview of the program is complete, it should be possible to move easily to the next step in the evaluation process and define more concretely the program's activities and participants.

CHAPTER FOUR

STEP 2: SPECIFY PARTICIPANT GROUPS TO EVALUATE

Introduction

To implement a program effectively it is necessary to know rather precisely: what *activities* the program will conduct; who will participate in these activities; and what the effects of participation are expected to be. To evaluate the effects of a program requires determining: what the program actually did; to whom it was done; and the results of having done it. In preparing the program overview, program plans or expectations were stated in fairly general terms. The purpose of this chapter is to suggest an approach to make the information in the overview considerably more specific as a base for further analysis.

The process and the prototype table presented below is designed to facilitate the necessary specification of data on objectives and activities as they relate to particular participants. In addition to facilitating the evaluation process, the information generated can also be used as a:

convenient reference summary which outlines program objectives, activities, and participants;

program management tool by focusing attention on the central aspects of the undertaking;

basis for formulating precise evaluation questions; and

basis for comparing different career education efforts with one another.

The Program Objectives/Participant Group Table

The product of this step in the evaluation process should be a table (or set of tables) which shows the relationship between program objectives, activities, and particular participants, whose performance will be analyzed further in the evaluation process. The end-product of this table is the identification, for further analysis, of groups of participants who participated in the same activities to achieve the same objectives. A prototype form providing this information is presented below in Table 4.1.

TABLE 4.1
PROGRAM OBJECTIVES/PARTICIPANT GROUP TABLE

(1) Program Objectives	(2) Grade Level	(3) Activities	(4) Program Participants	(5) Participant Group (Number)

A suggested process for completing the table is presented in the sections below. It should be noted that economies in the process described may be possible in some situations and, of course, should be pursued. Whatever the process and format followed, the time devoted to preparing a table which relates objectives, activities, and specific groups of program participants will be well worth the effort.

Suggested Process

1. List Program Objectives (column 1)

The major objectives of the program should be entered in column 1. As with the overview, only those objectives toward which the program is actually working should be listed. That is, do not list all of the objectives which everyone hopes might result from the program but toward which no real effort is being expended. To facilitate completion of the other columns of the table, from one half to a full page per objective is recommended.

Assuming that a reasonably careful job was done in preparing the overview (i.e., that objectives were phrased in unambiguous terms), these objectives may be recorded in column 1 of the table. If evaluation data is desired on more objectives than cited in the overview, then these additional objectives should also be recorded in column 1.

At this point in the process, it is suggested that objectives be stated in relatively broad, but unambiguous terms (i.e., as in the overview). While more specific behavioral objectives will be needed eventually, generally they can best be developed when formulating the evaluation questions and finalizing the evaluation design (Steps 3 and 8), rather than at this stage in the process.

2. List Grade Level for Each Objective (column 2)

The grade level(s) associated with each objective listed in column 1 should be entered in column 2. At least one line for each grade level should be allowed so that there is space for notations in the remaining columns of the table; as will be clear later, several lines for some grade levels may be necessary.

While this column may not be applicable for all objectives, it is not solely limited to those relating to student outcomes. For example, if an objective relates to teachers of the sixth grade, or to changes in school curriculum affecting the eighth and ninth grades, the applicable grade levels should be noted. A similar pattern should be followed with respect to parents. On the other hand, if the objective relates to changes in broad areas of school policy or to the relationship between the business community and the school, then it probably will not apply.

3. List Program Activities by Objective and Grade (column 3)

The major program activities which are expected to contribute to achieving each objective in each grade should be listed in column 3. If the same activities are associated with several grades (e.g., 4, 5, and 6), this may be noted on the table, thus saving needless repetition. Column 3 should be completed even when there is no entry in column 2 (i.e., the objective did not relate to a specific grade level).

It should be noted throughout the evaluation process that a career education activity is by definition something that is different from and/or additional to other activities provided by the school system. If field trips, for example, are cited as an activity, they should be of a particular type defined in terms of purpose, type of advanced preparation, or some other factor which distinguishes career education field trips from others. For instance, visits to art galleries for cultural enrichment would not be appropriate for inclusion, but trips to art galleries for exploration of career opportunities in the arts might be considered suitable.

The level of detail appropriate for this column is a matter for local judgment. In deciding the amount of detail, the following two questions should be considered at each grade level:

Is the activity expected to make a discernible difference in achieving the program objectives (or is it one of a number of things which can better be grouped together under a broader, but still meaningful, heading)?

Is the activity something the program wants singled out for investigation?

If the answers to these questions lead to the conclusion that several rather specific activities can be grouped together, then this should be done. For example, it might be concluded that those teachers who talk about careers regularly during class periods, have several guest speakers to discuss their jobs, and who regularly devote other class time to career-related television shows are providing their students with essentially the same exposure in pursuing an occupational awareness objective as those who talk about careers, devote time to television, but take their children on several field trips to learn about various jobs in lieu of having guest speakers. In such a case, the two overlapping sets of techniques might be grouped and considered simply one broad activity, i.e., classroom infusion of occupational information. On the other hand, in most cases it would not be appropriate to consider the use of special curriculum units, the regular exposure to the television series, or a large number of special field trips as sub-activities or techniques which can be combined unless all teachers are doing all of them. As a further illustration, if the cost-effectiveness of field trips had been questioned, it would be desirable to list field trips as a separate activity.

As indicated, the distinction between a technique and an important activity is one of local judgment. In general, however, it is best at this point in the evaluation process to err on the side of preparing a relatively long list of activities rather than to risk loss of detail which may be of potential importance.

4. List Program Participants by Activity (column 4)

Once the list of activities is complete, the next task is to identify program participants who are (or are expected to be) engaged in each activity, by grade level and objective. In some situations this will be a relatively straightforward operation and in others it may be rather difficult. Without knowing who participated in the program and how, it is impossible to determine accurately if the program had any effect.

While it is necessary to be able to identify each participant by name for further analysis, all of the names need not be listed in the table for each activity. Rather, it is more feasible and quite sufficient to group participants by teacher, course, school, or some other unambiguous group designation. For example, if the entry in the grade level column were "1st grade," and the entry in the program activity column were "classroom discussion by teachers," then the entry in column 4 might be: "all students in schools A and B," "the students in Mrs. Green's class in school C."

For the lower grade students and school personnel (teachers, counselors, etc.), this process typically is rather straightforward. At the upper grade levels where students tend to have more individualized experiences, the task may be a bit more difficult. In these grade levels it may be necessary to develop simple codes or nomenclature to accommodate small groups of students.

5. List Participant Groups by Objective (column 5)

In the final column of the table, groups and numbers of participants at each grade level who have engaged in the same set of activities related to a common objective should be recorded. Often the entry in this column may be the same as in column 4; this will not always be the case. Where column 4's focus is on individual activities, column 5's focus is on the relationship of participants to sets or clusters of activities.

The result of this effort will be a list of discrete groups of participants defined in terms of how they participate in the career education program. These groups can then be used as the basis for drawing samples and performing the other tasks which permit assessing the effectiveness of various aspects of the program. As with identifying participants, completing this column may be quite simple in some situations and complex in others. Essentially, it consists of inspecting the information in the other columns and recording in column 5 mutually exclusive groups of participants. Care should be taken to insure that the same participant is not included in more than one group for a single objective and activity. Table 4.2 on the following page is an illustration of a completed Program Objectives/Participant Group Table for a program at the elementary grade levels.

TABLE 4.2
ILLUSTRATION
PROGRAM OBJECTIVES/PARTICIPANT GROUP TABLE
Elementary Grades

(1) Program Objective	(2) Grade Level	(3) Program Activities	(4) Program Participants	(5) Participant Group (number)
1. Elementary students in schools A, B, and C will demonstrate increased knowledge of occupations and their entry requirements.	1	Teacher discussion	All in schools A & B; Mrs. Green's class in C	All in sch. A & B; Mrs. Green's class in C (158)
	2	Teacher discussion; TV series	All in schs A, B, & C	All in sch. A, B, &C (223)
	3	Teacher discussion	All in A & B	1. All in A (71)
		Visitors/trips	All in A	2. All in B (69)
	4	Infusion*	All in A & B	1. All in A (75)
		Curriculum units	All in A	2. All in B (72)
	5	Infusion	All in A & B	1. Mr. Gray's & Mrs. Red's class in A (47)
		Curriculum units	Mr. Gray & Mrs. Red in school A	2. Ms. Purple's class in A & all in B** (126)
	6	Infusion	All in A, B, & C	1. 35 observation site students in A (35)
		Observation sites	35 students in A	2. Other students (55) in A and all in B & C (180)
2. Elementary students in schools A, B, and C will demonstrate increased understanding of their own interests, aptitudes, and abilities.	1	Teacher-led class discussion	All in A & B; Mrs. Green's class in C	All in A & B; Mrs. Green's class in C (158)
	2	Curriculum units	All in A, B, & C	All in A, B, & C (223)
	3	Curriculum units	All in A, B, & C	All in A, B, & C (215)
	4	Curriculum units	All in A, B, & C	1. All in A & B (147)
		Small groups with counselor	All in C	2. All in C (76)
	5	Teacher led class discussion	Mr. Gray & Mrs. Red in A	1. Mr. Gray & Mrs. Red in A (47)
		Groups with counselor	All in C	2. All in C (65)
	6	Teacher led class discussion	35 students in A	1. 35 observation site students in A (35)
		Groups with counselor	All in C	2. All in C (70)
3. Elementary school teachers and the guidance counselor (School C) will utilize techniques, materials and curriculum provided by the program.	1	4 workshops; assistance from elem. coordinator	All teachers in A, B, & C	All teachers, A, B, C (9)
	2	4 workshops; assistance from elem. coordinator	All teachers in A, B, & C	All teachers, A, B, C (9)
	3	4 workshops; assistance from elem. coordinator	All teachers in A, B, & C	All teachers, A, B, C (9)
	4	4 workshops; assistance from elem. coordinator	All teachers, A, B, C; guidance counselor in C	All teachers, A, B, C; guidance counselor, C (10)
	5	4 workshops; assistance from elem. coordinator	All teachers in A, B, & C	All teachers, A, B, C (9)
	6	4 workshops; assistance from elem. coordinator	All teachers in A, B, & C	All teachers, A, B, C (9)

* Teacher discussion, visitors and trips, films, simulation.

** Assume there are three classrooms in A: Mr. Gray, Mrs. Red, and Ms. Purple.

In the illustration above there are two program objectives relating to elementary school students in grades 1-6 and one relating to teachers and a guidance counselor; this is shown in columns 1 and 2. From columns 3 and 4 it is apparent that for both student objectives the program provides several different activities, and the type of student involvement varies in the different schools and by grade level. With respect to teachers and the counselor, on the other hand, all are involved in essentially one general activity provided by the program.

In the illustration many, but not all, of the entries in column 5 are the same as those in column 4. Where all participants in a grade level engage in the same activity, then columns 4 and 5 are the same. This is the case for the first and second grade students. Where not all participants in a grade level engage in the same activities, then the entries in the two columns will differ. For example, in the case of the third grade with respect to the first objective, the students in school A engaged in both activities while the students in school B only engaged in special classroom discussions with their teachers (i.e., they did not have visitors or go on field trips). Because of this difference in the way they participated in the program, they were recorded as two separate participant groups. This was true of the fourth, fifth, and sixth grades as well. (It should be noted that accurate identification of participant groups is of critical importance in evaluation. Thus, in grade 6, Objective 1, 35 school A students form a group for evaluation purposes and the rest of school A joins with B and C to form another group for evaluation purposes. This distinction may seem rather fine, but will permit the project to assess the results of each experience or set of activities.)

With the completion of column 5, the relationship between specific groups of participants, the nature of their participation, and the outcome or result desired by the program is displayed. As indicated previously, a completed Program Objectives/Participant Group Table serves as a detailed summary of the program which can be used for a variety of purposes. One such purpose is step 3 of the evaluation process - - selecting evaluation questions - - discussed in the chapter which follows.

CHAPTER FIVE
FORMULATE EVALUATION QUESTIONS

Introduction

A comprehensive evaluation of a career education program may be expected to answer at least three types of questions. These are:

1. How do the experiences (i.e., activities, treatments, techniques, etc.) for program participants differ (a) among types of participants and/or (b) from non-participants?

2. How do the outcomes (products) for participants differ (a) among types of participants and/or (b) from non-participants?

3. How do the experiences relate to outcomes (in instances where there are outcome differences)?

In addition, the evaluation should answer questions relating to program management. Answers to questions in the management area are particularly important for explaining why activities did or did not occur as planned, and often hold the key to improving program operations. In addition, they will assist managers in dissemination of results and in future planning. They will also provide information necessary to school districts interested in replicating or adapting elements from the program. Additionally, answers to other questions in the management area may be required by parties not directly involved in the operation of the program.

Often the initial number of potentially important questions must be reduced to manageable proportions. Frequently, insufficient time is devoted by program staff to deciding which questions should be addressed and a hurried, more or less capricious selection is made. At other times selection is turned over to persons outside the program such as evaluators or consultants. In either case, the result is likely to be a set of questions that produces answers of little value and may even inject needless irritants into the school system. It is essential that the selections be made by program staff to assure that the questions cover the most important areas of interest and can be handled with available resources within the level of tolerance of the system.

In completing Step 2 of the evaluation plan (specifying participant groups), the basis for developing questions pertaining to the achievement of student and other program outcomes was identified. In addition, the base was laid for formulating questions regarding the extent to which related activities were provided. Step 2 should, therefore, serve as one important starting point for formulating evaluation questions. Other bases for identifying questions and some of the considerations which may be helpful in selecting the more important questions from among the large number possible are discussed below.

Formulating the Questions

It should be noted that the scope of the evaluation may vary in terms of the number of questions it addresses. That is, a school district may wish to answer a large number of questions in each of the areas discussed above, or only a select few. Generally, the scope of the evaluation depends upon the nature of the program, the purpose of the evaluation, and such limitations as time, funding, resources, and the "quality" of one's data. Regardless of the scope of the evaluation, choices should be made in a systematic fashion, and a series of questions concerning priorities should be addressed. These include:

> What institutions and individuals (interested parties) are most concerned with the evaluation and its results? What priorities should be established among them?
>
> What would the interested parties like to know and which questions are the most important?
>
> Why do these parties want the questions addressed (e.g., legal obligation, basis for making specific types of decisions, curiosity, etc.), and how precisely need they be answered?
>
> What will it take to answer these questions (e.g., answers to a series of more specific questions, a two or three year research study, etc.)?
>
> What are the major constraints on the evaluation (e.g., cost, limits on testing students and questioning school staff, etc.)?

> Which questions are obviously impossible to answer given the program's objectives, activities, and constraints on the evaluation?
>
> Of the remaining questions, what is their order of importance to the program staff, in terms of legal and political realities and the potential for contributing to program improvement?

Ideally, the answering of these questions should be undertaken by a small committee of people intensely interested in the program.

In the course of preparing an initial list of evaluation questions, it is important to review grants, contracts, school district policy, and other legal and quasi-legal documents which bear on the program as well as the Program Objectives/Participant Group Table (Step 2). If the career education program is supported by federal, state, or private funds, several evaluation questions may be required by the funding source. Similarly, the school board or others in the community may virtually mandate the inclusion of some questions.

The issue of identifying the audience(s) for the evaluation and their needs is of crucial importance and is often overlooked. If the evaluation is to be of practical benefit, it is important to clarify at the outset for whom each question is directed and why. Questions designed to provide information to policy makers removed from operational responsibility for the program generally will differ in content from those designed to provide information for the program director, program staff, or teachers. Decisions concerning how much information is to be collected, of what type, and at what level of detail will all depend on the type of audience to which the answer is addressed. Experience suggests that the failure to go through this process of identifying the relevant audiences and their needs often results in evaluations of limited practical worth.

Ultimately, the ability to answer evaluation questions will depend on such factors as the availability and reliability of data sources, instruments, or other measuring devices, the timing of the evaluation, and, more generally, the questions of appropriate evaluation design. During the early stages of formulating the evaluation questions, however, it is suggested that these issues receive only minor consideration. Unless the program committee is unusually experienced in evaluation, it is best to list in order of priority what the group considers to be the most important questions, and later formally to record any reasons which preclude certain questions being addressed. By following this approach, the program manager is provided with a sound and ready response to inquiries from interested parties who did not have their questions answered.

Once the initial set of questions has been listed, it is useful for the committee to review the questions from two points of view: first, the appropriateness of the questions to the program, and second, the implications of how they are phrased for the evaluation design.

Appropriateness

As indicated earlier, career education staff may often be required to answer a number of questions not of their own choosing. These may have been raised by the school board, other influential persons in the community, funding sources, or some other audience to whom the program feels obliged to respond. Sometimes these questions are quite germane to the program, but at other times they are not.

Given the intensity of interest in career education, it is not uncommon for programs to be asked to respond to questions which are completely outside the scope of local operations. For example, a program may be operating only at the elementary school level, yet a question is raised regarding outcomes at the senior high school grades. Assuming the question must be answered, the only appropriate response is a simple statement that the program is not intended to affect senior high school students and therefore no impact is anticipated. In such a case, there is no reason to devote evaluation resources toward answering this question. If conditions demand that the question must be acknowledged in the evaluation report, a sentence or two stating why it was not investigated in depth (perhaps also stating why the program did not focus on the subject of the question) should be sufficient. Knowing that the question was not germane may be quite important to those who raised it.

It is often the case, however, that questions which at first appear to be outside the scope of a program may be appropriately addressed. For example, a program may be asked to assess its impact on student reading achievement even though its statement of objectives does not address this area. Similarly, a program designed to deal exclusively with teachers may be asked to assess its impact on one or more areas of student learning. In such a situation, it is first necessary to conclude whether it is at all reasonable to believe that the program may have had an impact relating to such questions. Second, it should be decided, on the basis of the position of the party raising the question or a newly found local interest, whether the question should be addressed. If it is to be addressed, it should be made clear in reporting the answer that the questions did not pertain to an area of primary program concern. If it is not to be addressed, the need for a brief explanation for its omission should be considered.

Initial Design Implications

The basic design of the evaluation and the form of the questions are tightly intertwined. Thus, once the appropriateness of each question is determined and priorities among questions set, it is useful to review the list for the design implications implicit in the phrasing of the questions. For example, the use of relative terms in a question may necessitate the use of pre-post measures and/or comparison groups. Similarly, the standard or criteria by which program success or failure is to be judged may or may not be included. It is quite possible that a change in wording will make an impossible question possible to answer.

While it is essential that precise measurable questions be defined prior to implementing the evaluation, at this point in the process some interaction between the evaluation questions and the results of subsequent steps is to be expected. Thus, unless the evaluation design is to govern many details of program operations, at this stage it is generally advisable to permit some flexibility in the way questions are phrased.

Once this review of the initial set of questions is complete, it is appropriate to proceed to step 4 of the evaluation process - - specifying data sources.

CHAPTER SIX

IDENTIFY DATA SOURCES

Introduction

Once the list of evaluation questions has been prepared, the next step is to identify the sources of information which will provide answers to each question. In practice, the data sources utilized to answer each question will depend both on the location of the data and on the availability of the instruments necessary for collecting this data. Since the possible sources of information limit the type of instrument(s) selected or developed, sources should be identified before consideration is given to development of the instruments.

In some situations, identifying the appropriate source of information to answer a question will be clear cut, in other situations more complex. In part, this will depend on the question being asked and in part on local conditions. For example, in some districts, it will be most appropriate to obtain information on relevant areas of student learning directly from students participating and not participating in the program through tests and questionnaires; in other cases, some or all of the information needed to answer a particular question will be available in school files and there will be no need to approach students directly.

In considering possible sources of information, five basic questions should be addressed:

> What source has the information needed to answer the question?
>
> How accessible is the information (i.e., are there limits to the amount of information which can be obtained or on the ways it can be obtained)?
>
> How valid and reliable is the information (i.e., will the source consistently provide accurate information or will it be erratic or biased)?
>
> Are there other sources of information?
>
> Which is the most efficient source to use?

Identifying Possible Sources

Typically, the same group that formulated the questions should be used to identify potential sources. It is also quite helpful at this point to involve school principals and members of the central staff of the district who may be aware of sources of useful information already available or which is to be collected for other purposes during the year. Since the selection of sources so directly impinges on the areas of instrumentation, sampling and design, it is suggested that the persons responsible for these aspects of the evaluation also be included at this point in the process.

In seeking possible sources of information, it is often helpful to think of the following categories: direct vs indirect sources and primary vs secondary sources. A systematic procedure for considering and choosing adequate sources in each category requires little time or effort. Moreover, it may produce results which lead to major efficiencies in implementing the evaluation.

Direct sources are individuals who are the subject of the question. For example, in a question concerning the extent to which students in the program learned more than students who did not participate, the two groups of students would be "direct" sources of information. Indirect sources are individuals possessing knowledge about the subject of the question. For example, if the question concerns student learning, then their teachers would be an indirect source.

Both direct and indirect sources can be subdivided into primary and secondary sources. A primary source is an original document or a first hand account or recording of events; a secondary source provides material copied or compiled from primary or other secondary sources.

If the career education program or the school system in which it is operating is relatively complex, it is sometimes helpful to formalize the search for alternative data sources. One possible approach is to use the following simple matrix keyed to each evaluation question and to enter possible sources in each cell as appropriate.

Data Source Matrix

	Direct	Indirect
Primary		
Secondary		

Even in rather simple situations, a systematic, mental review of possibilities in each of these categories has led to finding sources of information which greatly facilitated the evaluation.

Selecting Sources

The final selection of the data source(s) for each evaluation question should normally await consideration of instrumentation, sampling, and design. However, in cases where there is only one possible source of available, valid information, then, of course, there is no alternative.

In rare cases, no source of valid information may be identified. In this event, it will be necessary to modify or eliminate certain evaluation questions.

Where several possible sources have been identified, however, it is useful before proceeding to the next steps in the evaluation to identify the one or two preferred data sources for each question. In selecting from among several possibilities, consideration should be given to the validity and accessibility of the information. It is also important to consider the burden on individuals of the combined effect of using them as the source for several questions. For example, there may be school district limits on the amount of time students can be tested or teachers questioned. Excessive demands on the time of teachers, counselors, and other school staff may also adversely affect the quality of the data they provide.

Based on the identification of one or more sources of accessible, valid information for each question, the next step in the evaluation process is to select or develop appropriate means of securing the information. This is the subject of the two chapters which follow.

CHAPTER SEVEN

SELECT OR DEVELOP INSTRUMENTS OR OTHER MEASURES

Introduction

As discussed in Chapter Five, evaluation questions should address the areas of program objectives (i.e., outcomes), program activities, and program management. Once the set of evaluation questions has been phrased and potential sources of related data identified, the next step in the evaluation process is to select or to start developing instruments for recording data needed to answer the questions.

Depending on the nature of the questions posed and the data source(s) available, one or more of a wide variety of types of instruments may be most appropriate. These include:

- standardized tests or questionnaires;
- specially developed tests or questionnaires;
- observation scales;
- interview schedules;
- document review summary forms; and
- staff reporting forms.

In selecting from among and within these categories, important factors to consider are the instrument's substantive relevance to the question, the validity and reliability of information obtained, and the cost of securing sufficient copies and analyzing the information.

Perhaps most important from a program management point of view are the considerations of relevance and ease. Because of the particular importance of these two elements, before an instrument is used in the program it should be reviewed carefully by the program manager to insure that the specific questions relate to program objectives or activities and that its use will not cause major administrative problems (i.e., violation of privacy laws, excessive complaints from parents, staff, etc.).

Major Types of Instruments

In the following sections we discuss each of these six types of instruments and some of the main considerations in their selection and use.

1. Standardized Tests or Questionnaires

A number of standardized tests and questionnaires for students are identified and described in the Appendix of this handbook. They have been screened and evaluated by a special panel of experts against the following criteria:

- the test could be administered on a group basis;
- student responses required little or no writing;
- the instrument's objectives could be related to one or more of the objectives of career education;
- psychometric (i.e., reliability and validity) data were available for review and judged adequate; and
- each instrument could be obtained by local school districts for use in program evaluation.

In making decisions on the advisability of using standardized instruments or developing instruments locally, the program manager should give particular attention to the criteria discussed in the introduction to this chapter. In the case of the tests described in the Appendix, a panel of experts has identified areas of substantive relevance, and assessed the tests' validity and reliability. For each test recommended by the panel, the Appendix provides basic information regarding relevance, administration, and cost.

The tests described in the Appendix will meet many needs in the evaluation of career education. Clearly, however, the listed tests will not meet all of the needs. For example, standardized tests in mathematics, reading, and other skill areas traditionally used on a regular basis by school districts and which might be considered responsive to some evaluation questions were not included in the review. Nor were standardized instruments which relate to specific, commercially available curriculum packages or those which address specific vocational skills (e.g., electronics, typing, etc.). Reputable sources of information about standardized instruments relevant to areas not directly covered in the Appendix include:

- The Mental Measurements Yearbook; Buros, Oscar Krisen; The Gryphon Press. Highland Park, New Jersey.

- Center for the Study of Evaluation; UCLA Graduate School of Education; Los Angeles, California.

- Educational Testing Service: Test Collection; Princeton, New Jersey.

- Educational Resources Information Center (ERIC) Clearinghouse for Career Education and for Tests, Measurements, and Evaluation.

- State Departments of Education.

Information may also be available from the Office of Career Education of the U. S. Office of Education and the Education and Work Program of the National Institute of Education.

In addition, test publishers themselves are valuable sources of information. Evaluators and program staff should feel free to contact test publishers or authors and request both general and quite specific information. Often the test developer can provide frank information regarding the relevance of an instrument or he may be able to suggest or provide special scoring procedures which enable use of an instrument in the evaluation when it otherwise would not be directly relevant.

When seeking standardized instruments suitable for teachers, counselors, parents, and others who may be important to the evaluation, the sources cited above should be considered. Teacher attitude and knowledge scales with potential relevance to career education programs are becoming increasingly available. Similarly, commercially available instruments exist which may be appropriate for parent and community surveys.

In general, when criteria of relevance, ease of administration, and cost are met in a particular setting, the use of standardized instruments is preferable to use of tests developed locally. Typically, greater confidence can be placed in the technical characteristics of such instruments and they greatly facilitate comparisons across programs.

2. Specially Developed Tests or Questionnaires

a. Tests

For information not available through standardized instruments, it may be necessary to consider the development of instruments locally. If it is decided that the need for one or more "home-made" tests justifies the time and costs involved in development, there are several important technical factors which must be considered. Some of these factors apply to each item to be constructed; others apply to the "test" as a whole. In either case, the major considerations pertain to the need to assure maximum validity and reliability of the responses, and ultimately of the score or scores which will be generated by the test.

Test construction, however, is a highly specialized field. Before embarking on a program of locally developed tests, it would seem essential to obtain the services of a staff member or consultant with the necessary expertise. Just a few of the technical issues that will require attention are identified in the discussion which follows.

The starting point for maximizing validity is a clear delineation of the subject matter, or outcome area, which the test is supposed to measure, followed by a detailed listing of specific content areas to be covered within the outcome area. The test items would then be constructed around the specific facts or concepts which represent the more important teaching objectives.

Having specified the content areas of the test items, a decision is needed concerning the format to be used for the items themselves. There is usually a choice among several alternatives, such as True/False, multiple choice, Agree/Disagree, and matching, among others. It is more difficult to construct items in some formats than in others, but the effort may be worthwhile. For example, multiple choice items are more difficult to construct than true/false items, but the results are apt to be more stable and more valid since there is a 50/50 chance of getting a true/false item right simply by chance. Where there is a high probability of being correct on the basis of chance alone, test makers generally construct a relatively large number of items as a way of compensating for chance effects and insuring adequate spread in the scores achieved by the respondents. In addition to the ease or difficulty of constructing the items, the choice of format should also be based upon such factors as: the suitability of the format for the intended respondents; the length of time to administer the test using each format; and ease of scoring.

Prior to devoting the resources necessary to develop adequate test items locally, it is advisable to review the items on standardized tests and those available through national research organizations. One such potentially useful source is the test item pool in the area of career and occupational development available through the National Assessment of Educational Progress.[2]

In the course of selecting or constructing the items, great care must be taken to insure that there is no ambiguity in the wording or response alternatives and to insure that the vocabulary and sentence structure are within the capabilities of the potential respondents. The language of the item, as well as the test formats, must be so designed as to insure that incorrect responses occur only because the student does not know the correct answer, rather than as a result of difficulty in understanding the item or confusion with regard to the test directions. Both of these factors may be checked by means of a small-scale pretest of the instrument involving 20 to 30 respondents, providing that the group includes as wide a range of ages and abilities as the group of intended respondents.

In the course of revising items to insure that the language is clear and appropriate, it is easy to lose sight of the specific knowledge which the item was intended to assess; often the item may be easy to understand but has lost direct relevance to the outcome area. This fairly common occurrence may be minimized by having cognizant teachers review each item for appropriateness of both the substance or content and specific language and format. They should also review the test as a whole for such defects as repetition of content areas, inadvertently providing the answer to one item in another item, under-sampling or over-sampling of the priority facts and concepts involved, and appropriateness of the level of difficulty.

---

[2] National Assessment of Educational Progress; 1860 Lincoln, Suite 300, Denver, Colorado 80203.

In summary, the development of adequate tests is a complex, time-consuming undertaking. While there clearly is a need for additional instruments which adequately address student outcomes associated with career education, development of tests should be attempted only with great care.

b. Questionnaires

Questionnaires may need to be developed to obtain some information not available in pre-existing documents or which cannot be obtained through tests. It is likely that separate questionnaires will be needed for teachers, counselors, students, administrative staff, and others who provide or observe various activities. As used here, a questionnaire is defined as an instrument which is given to an individual with the understanding that the recipient will carefully answer in writing the questions asked. Such questionnaires can be used to gather information on both activities and outcomes.

Ideally, a questionnaire would go through the same rigorous process of development as a test. In practice, however, this is not usually possible. Nevertheless, in the context of evaluating a career education program in a school district, the questionnaire will often be the most feasible method of accumulating answers to some evaluation questions. Below are some guidelines for questionnaire development. Since some of the technical problems in questionnaire preparation, though of a lower order of difficulty, parallel those in test construction, the services of a staff member or consultant with the necessary skills will be important to the development of such instruments.

In developing a questionnaire there are several basic rules which should be followed carefully. While they sound simple and have been repeated in the literature for years, they are violated frequently enough to warrant repetition.

The primary rules are as follows:

The questions should be simple and precise. Perhaps the most common difficulty in question construction is the combination into one question of what should be two or more questions. For example, if program personnel wish to ascertain how many of the students in a ninth grade English class that went on a particular field trip wrote an essay about a particular field trip, several short questions are necessary. It is not enough to ask: "Did you write an essay on your field trip?" A negative answer, of course, could mean that he did not go on the field trip or that he went on the field trip but did not write an essay. The first question should be: "Did you go on a field trip to X?"

The questions should be understandable. The words and concepts should be readily understood by the reader. For student questionnaires that means vocabulary appropriate to the grade level. It also means avoidance of ambiguous phrases. For example, if the need is to know how much time a counselor spent in career counseling, it is crucial that all recipients use the same definition of the concept "career counseling." If there is doubt, the question should be broken into various components.

The criteria for providing the response should be clear. At times there are two or more perfectly logical ways to answer a question. For example, depending on precise phrasing, three "correct" answers might be given to a question relating to the number of students seen by a counselor who had interviewed five students, three of them twice: Total interviews - - 8; different students interviewed - - 5; students interviewed more than once - - 3. In designing the questionnaire the first decision should be to determine precisely what information is needed and then to phrase the question to get that information.

The response should require as few words from the respondent as possible. Ideally, questions and possible answers can be so well thought out that the respondent will simply need to place a check mark in a designated square. This reduces responding and interpretation time. Often, however, this is not possible and open-ended questions are necessary. Where this is the case, tentative categories of responses should be prepared. In some cases it may be concluded in advance that there will be so little commonality in responses that an item should be excluded.

Catch-all response items such as "other" should be avoided where possible. The existence of such items suggests that the question was not sufficiently thought out or that the information desired was not clear. Often such items produce no helpful information. There is also the danger that significant data will be lost because of the unspecific response.

In essence, most questionnaires are collections of single questions put together such that they are all asked on the same form. The construction of the items and the order in which they appear, then, are the two most important factors. In item construction the preceding rules should be carefully followed. It may be useful to review questionnaires from other projects to identify items which were used successfully. In using items developed elsewhere, however, consideration should be given to the unique characteristics of the local respondents, as the interpretation given to certain words or phrases may vary from location to location. Questionnaires used in national-scale undertakings are potential sources of items which overcome problems of local terminology.

The primary consideration in ordering the items or sets of items is: (a) their influence on responses to subsequent questions; and (b) the attitude engendered in the respondents. In the first instance, attention must be given to the possibility that the answer to one question will dictate somehow the response to other questions when this is not desired. In the second, care must be taken to insure that the time necessary to respond to the first item will not lead the respondent to become resentful or angry and discontinue completing the questionnaire, or provide erroneous answers.

3. Observation Scales

Observation scales are primarily used for quantitative assessment of behavior(s) of teachers and students in order to obtain information which is generally considered qualitative in nature (e.g., classroom atmosphere, teacher attitudes, etc.). Properly designed, this instrument identifies key points to be observed and thereby provides a basis for consistency of observation in different settings. The instrument may also provide a scale of frequency, intensity, or time alloted for certain types of activity to introduce a measure of objectivity into what might otherwise be areas of subjective judgment.

The advantage of such scales is that they permit an objective outsider to observe administration of a critical activity; for example, the behavior of the teacher and pupils and their interaction. These observations may verify teacher reporting forms and/or permit identification of facets of classroom activities which are not acquired through other data collection techniques. In a similar manner, the observation scale used in observing visits to industrial plants may shed light on, or help reinforce, other measures of student learning during field visits.

The difficulties with this approach are two-fold. First, it is never possible to obtain as many observations as might be desired. Implementation is time-consuming in that a third party must observe the treatment at least once and ideally several times. Second, when more than one observer is used, it is difficult, if not impossible, to be certain that all are using the same criteria in making their judgments. Furthermore, observation of an event is subject to influences beyond the control of the observer and descriptions of the same event from the same person may change over time.

If this approach is utilized, it is important to develop observation scales which require as little subjective judgment as possible. Thus, it must be clear to the observer what to look for and record. As with questionnaires, the more often the observer can respond by the use of a check mark the better. Second, it is important if several observers are used that an assessment of interrater reliability be made. This simply means that several observers should view the same event and record their findings. The extent to which their reports are the same is the extent to which reliability among raters has been achieved.

4. Interview Schedules

The interview schedule or interview guide is an instrument to insure uniformity of items to be obtained orally from one or more sources and recorded in standard format. The interview schedule may be used to supplement any of the other methods of data collection or to verify reported information or assist in its interpretation. For some purposes it may be the only source of information.

Such schedules are administered in a face-to-face interview or by telephone. The format may be quite tight, permitting little or no deviation from what is essentially a script for the interviewer, essentially a questionnaire to be answered orally. Or it may be quite loose, permitting a conversational atmosphere, with a high probability that the order of the questions to be asked will not be as they appear on the guide and that the answers may not be directly responsive to the questions.

The appropriate format to be adopted will depend on the use to which the information is to be put and the number and status of the interviewees. If there is a single interviewee (e.g., the school superintendent), then both methodological and pragmatic considerations suggest the looser format. There is no need for a precise format, as the answers will not be compared directly with those of similar individuals, and as a practical matter the superintendent is likely to resist a rigid interview format. On the other hand, if the purpose is to elicit standard responses from a group of teachers or students, then a more structured approach may be necessary to facilitate tabulation and analysis of the interview results.

As the sole source of information about activities or outcomes, there are three situations where this approach to data collection may be preferred:

> When the respondents are unlikely to respond to any other method. This may occur when information from employers providing training to high school students is necessary. Often there is greater possibility of obtaining a brief interview with an employer who supervises a student in a cooperative education program than there is of securing a written response to a questionnaire.
>
> When the respondents are expected to have difficulty with a written form. This may be due to the respondents' fluency in written English or the inherent ambiguity of some terms associated with the educational program. In both cases it may be judged essential that an interviewer have the opportunity to be certain that the respondent fully understands the questions asked.
>
> When the purpose is to explore an area with the respondent. At times, what is sought is the result of following a chain starting from a precisely formulated question which could be answered in a multiplicity of ways. It may be, for example, that when the list of program activities was developed it was impossible to define precisely what happened in some locations and it may be necessary to interview the teacher, several students, and employers to identify the specific nature of the activity. From the results of these open-ended interviews would emerge the needed description of the finer points of the activity category.

In constructing interview guides, the general rule for constructing other types of questionnaires should be followed.

5. Document Review Summary Forms

Data pertaining to both activity and outcome questions may be included in documents regularly prepared by the school district for other purposes. For example, in some school systems it may be possible to assess the extent to which a career information center is utilized by various groups of students by reviewing the records kept by the center staff or school librarian. If the center has been in operation for several years and if comparable records exist, it may be possible not only to assess the extent to which particular groups have used materials during the evaluation period, but also to assess the extent to which they had used materials in prior years. By the same token, it may be possible to obtain information on the number and type of field trips taken in some schools by a review of school transportation records. In many situations counselor records may also be a significant source of information.

As is true in the development of all instruments, the key in developing document reporting forms is to know precisely what information is desired. Presumably this will have been accomplished in the process of identifying treatments and formulating evaluation questions. It is of equal importance to have a thorough knowledge of the documents which are to be reviewed prior to developing the forms, because school files and other such documents are used primarily for other purposes and are rarely designed to provide precisely the information desired in a program evaluation. As a result, it is necessary to develop summary forms to obtain the maximum information available with the least amount of effort.

In essence, document review summary forms need to be:

- specific as to the information to be selected and reported;
- organized to make selection as simple as possible based on the format of the school records and reports; and
- arranged to permit ease of tabulation and analysis.

Beyond the documents themselves, it is important to know how the school's reporting system actually operates. It is not unusual, for example, for a cursory review of documents to suggest that they contain information which, upon closer analysis, is found to be internally inconsistent or incomplete. Further, the records may not report what they seem to be reporting. Analysis may show, for example, that what appears to be the number of students placed on jobs is in reality the number of referrals of students to potential employment. Another type of problem may be found when seeking the number of students in two or more activities: the records may appear to indicate that 200 students are involved in various program activities when in fact a group of 50 students is participating in four different activities. Further, it is not unusual for records to contain estimated figures which for most purposes are quite reasonable but may not reflect subtle but significant changes and so may prove to be misleading when used in assessing program outcomes.

In summary, while obtaining information from existing documents may at first seem to be the preferred approach on the grounds of ease of data collection and convenience to the school district, experience suggests that this is not always the case. Before concluding that this approach will be followed and finalizing the summary collection forms, it is desirable to develop the instruments based on the criteria cited above and then undertake a trial run to determine whether in fact this source will be as productive and economical as it would appear.

6. Staff Reporting Forms

Staff Reporting Forms can serve as the means of collecting data from original sources - - e.g., teachers, counselors - - primarily on matters of program implementation. They can call for information on a one-time or periodic basis or as a by-product of other activities or reports.

Given the advantages to program management that this approach provides, there are nevertheless certain constraints on extensive use, and it should therefore be used primarily as a supplementary data collection device.

First is the problem of precisely identifying the information which is to be reported and the development of forms which capture the information in a consistent manner across all staff members. This is the same problem confronting developers of all types of instruments, made more acute by the possibility that user irritation with the reporting system might affect their interest in the career education program.

Second, there is the problem of accurate reporting of information. Heavily burdened teachers and counselors may provide less accurate information than is needed. This problem can be met, in part, by careful design of instruments, stating questions simply and accurately, providing boxes and spaces for inserts and minimizing the amount of narrative. In addition, simple procedures can be developed to verify the reporting information. These might rely on comparing the self-reported information with school records concerning some of the same areas. For example, field trips reported could periodically be cross checked against school transportation records. Class observation is another method which could provide a periodic check on the results of the self-reporting forms. Finally, students or others involved in the process could periodically be asked questions which would provide verification. Whether the cross-checking offsets the value of use of staff reporting is a question that would require careful attention.

These instruments, primarily intended to develop data for program evaluation, can be helpful to teachers and administrators in monitoring their own progress in implementing their programs. If, for example, a plan calls for four field trips during the fall semester, interim reports will help in judging whether particular classes are on schedule or behind. Similarly, if a plan calls for certain outcomes by Thanksgiving, the instruments will help the teacher and principal examine whether they are on target without waiting for a final evaluation report which may be a year away. Or, they may suggest that for a variety of reasons the original plan was unrealistic and that new targets should be established.

The principal purpose of the Staff Reporting Forms is to get primary information from the original source which cannot be obtained in other ways, or to supplement, or corroborate, or verify information obtained from other sources.

The second use of these instruments is to provide a summary record at the end of the year of the activities which were administered to various groups of students. While this information might be obtained through an interview or questionnaire given to staff at the end of the year, it is likely that the information will be more precise and accurate if the staff is reporting these events on an ongoing basis.

The third and most significant constraint is the imposition such systems make on staff time. If the form requires considerable time for the staff member to complete, there may be political or accuracy costs to pay. At the one extreme, teachers may simply refuse to complete the forms and at the other they may complete them carelessly and inaccurately. It is important, therefore, to develop instruments which are quickly and easily completed and for which the staff can see a useful purpose. A number of career education projects have had success in utilizing this method in collecting a major portion of the treatment data and have developed forms which could be of use.

In developing staff reporting forms or procedures, most, if not all, staff objections and inherent difficulties mentioned above can be overcome if the reporting form is a natural by-product of normal action instead of something imposed in addition to normal procedures. If, for example, counselors write a simple referral sheet when referring students to jobs, it is better to have a carbon copy of that referral serve as the reporting form than to have a separate summary form for the counselor to fill out. Summation then becomes a routine task which can be performed by support staff on a regular basis. Similarly, if selected items can be incorporated in or added to regular reports required by the principal or the school district, there is no appreciable increase in the work load of the user and the data can again be tabulated periodically by support staff.

Summary

Regardless of the approach or combination of approaches used to obtain information regarding a specific evaluation question, it is clear that instrument selection and development require careful, time-consuming labor. In all cases precision is required with regard to the information sought and an understanding of the uses to which the information is to be put. Also required is an assessment of the practicalities of data collection such as ease of administration, and the time required. Most important, attention must be paid to the issue of obtaining valid and reliable responses; the various cautions, guidelines, and suggestions contained in this chapter are designed to help in this regard.

Another factor to be considered in the development of instruments is the ease of quantifying activity variables. Some will be relatively simple, such as the number of field trips or the number of different occupations represented by the outside visitors to a classroom. Other potentially influential factors, such as the emphasis placed by the teacher on career decision-making, may lend themselves only to very coarse quantitive categories; e.g., "much emphasis," "little emphasis," or "none at all." Regardless of the method used for quantification, or the precision of measurement which may be possible, it is essential that the instruments lend themselves to the quantification of each variable in order to provide a basis for assessing their relationships.

CHAPTER EIGHT

CONSIDER DESIGN AND ANALYSIS ALTERNATIVES

Introduction

After formulating the evaluation questions and making at least tentative decisions regarding data sources and instrumentation, the next step in the process is to consider evaluation design and analysis alternatives. Individuals with technical training and experience should be assigned prime responsibility for evaluation design and analysis, but since design decisions should reflect program needs and constraints, the program manager and staff should participate in the decision-making process and review the decisions made.

Evaluation designs, employed in an attempt to generate objective information concerning project effectiveness upon which to base management and policy decisions, are generally patterned after research or experimental designs, but may differ significantly from them. The most obvious difference is that they cannot always meet the rigorous requirements of experimental designs. Another difference frequently cited in the literature relates to the nature of the subject under investigation. In experimental sciences the researcher generally specifies the area of interest or investigation and generally is able to control most aspects of the situation. In evaluation, most aspects of the situation are outside the control of the evaluator and the area of inquiry is specified by the client.

The selection of specific techniques to be used in analyzing the data collected should be an integral part of preparing the evaluation design. Since different techniques are appropriate for different types of data, it is desirable to specify the preferable technique prior to developing a sampling plan. In other words, the sampling plan should be designed to fulfill the requirements of the preferred design and analysis approach.

The selection of the most appropriate evaluation design cannot be separated from the evaluation questions and the information available to answer them. In addition, the design selected is often closely related to the institutions or individuals interested in particular evaluation questions and the use they will make of the answers. Thus, it is not unusual for different designs to be used to answer different questions. For example, if an external funding source requires sound evidence of the impact of the program on reading scores in the elementary grades, then a rigorous research design is essential. However, if the question relates, for example, to whether or not a pre-specified level of accomplishment was attained (e.g., placing 80% of the graduating seniors in jobs or further training), then an alternative design will provide adequate information.

Generally speaking, it is preferable to select the design which permits answering each evaluation question with the greatest possible confidence. Since the most confidence can be placed in answers based on the research designs of controlled experiments, if it is necessary to "prove" something (e.g., that career education increases reading achievement scores), then true experiments (with control groups) are necessary. Although there are several "true" experimental designs which can be used, a major problem in the design of educational program evaluations occurs when the requirements of these "true" designs cannot be met. In a school setting, the most troublesome requirement is to randomly assign students to program activities (treatments).

While it is often not feasible to meet the requirements of true experimental designs, they should not be rejected automatically. When the evaluation and the implementation of the program are planned together, such designs can be used far more widely than is current practice. Even after a program has already started, "mini-experiments" keyed to specific but potentially important questions (i.e., the relative effect of one activity versus another) may be possible. Thus, while it may not be possible to answer most evaluation questions on the basis of true experimental designs, it may be possible to answer some of them.

When the requirements of the "true" research designs cannot be met, there are several alternatives. The section below includes a brief discussion of five designs often used in educational program evaluations. The five designs, beginning with a rigorous research design, cover a range of the most likely situations which will occur. While there may be less certainty in the answers obtained from the non-research design alternatives, generally they yield approximations which are quite useful from a management perspective. However, since there may be legitimate disagreement over how to interpret specific findings, it is important that the procedures followed be documented and understood by potential users of the information.

Design Alternatives

There are a variety of designs which may be appropriate in a program evaluation. The most appropriate design for a given evaluation question will depend on:

- the information available (i.e., sources and instruments);

- program constraints (e.g., time, cost, control over activities and participants, etc.); and

- the precision and confidence with which it must be answered (e.g., rigorous proof, quantitatively supported estimate, reasonably informed assessment).

In selecting from among design alternatives, program management as well as technical, analytical judgments must be made. In a school setting, compromises are almost always required between program-imposed constraints and the desire for precise and confident evaluation findings. While the evaluator may be expected to outline alternatives and to raise and respond to questions, striking the final balance between constraints and desires is the role of the program manager.

Presented below are several designs which experience indicates will be relevant for most career education programs. As suggested earlier, it is unlikely that any one design will be appropriate for all questions, and therefore more than one of those discussed may be used in a single program. There are also more complex designs, of course, which are appropriate in some situations. More detail regarding the designs below and several other alternatives may be found in texts on evaluation and research design.[3]

Posttest Comparison with Matched Groups

This is a research design which requires pretesting a large group of potential participants (at least twice as large as the desired number), pairing those tested in terms of pretest measures, and then randomly assigning one member of each pair to the participant group and the other to the comparison group. It is important that the posttest measure the same skills as the pretest. This design provides an accurate estimate of the impact of the program (or activities) on participants. At the same time, it is also a more difficult design to use in a school setting.

---

[3] Of particular interest to the manager may be: Donald T. Campbell and Julian C. Stanley, Experimental and Quasi-Experimental Designs for Research; Chicago: Rand McNally, 1963.

D. Horst, K. Tallmadge, and C. Wood, A Practical Guide to Measuring Project Impact on Student Achievement; Washington, DC; US Department of Health, Education, and Welfare, Office of Education, 1975.

Pretest-Posttest Control Group

This is a research design which requires: (a) random assignment, prior to treatment, either of individuals to participate and comparison groups or of program activities to some existing but essentially similar groups (e.g., classrooms) and not to others; and (b) pre- and posttesting the participants and comparison groups. It can provide an accurate assessment of the impact of the program on participants only if the pretest difference between participant and comparison groups are due to chance. Although it may be difficult to meet the requirements of this design on a large-scale basis (e.g., throughout the program or the school district), it may be quite feasible in many situations.

Norm Reference Design

This design requires use of standardized tests with pretest and posttest norms and careful adherence to the testing procedures followed by the test publisher in obtaining normative data; it does not require local comparison groups. It is particularly important to note that testing dates should correspond to those used in norming the test (e.g., October and May), and that the norm group should be as similar as possible to the participants. This design may provide a quantitatively supported estimate of the impact of the program on participants. Where standardized tests are available and proper testing procedures followed, this design can be implemented at considerably less cost and effort than a comparison group design.

Criterion Reference Design

This design is most appropriate for program objectives dealing with levels of performance (e.g., the number of materials developed, the extent to which teachers use a media center, the number of students placed in employment, or the number of courses added to the school curriculum), or verifiable changes in the school district (e.g., a school policy favoring career education, an agreement with the employment service, etc.). It provides an estimate of the extent to which a predetermined criterion has been achieved. It requires a careful specification of the criterion, and valid and reliable pre and post measures, but may not require a comparison group. Although useful in measuring how close performance has come to the criterion, unless it is combined with one of the previously cited comparison group designs, it does not establish that the performance was caused by the program.

Survey Design

This design can provide a description or a quantitatively supported estimate of various relevant characteristics of a group (e.g., knowledge, attitudes, social characteristics, etc.) at a particular point or several points in time. It requires carefully developed questionnaires which address these areas of interest. While it does not permit drawing the conclusion that program activities were responsible for the survey results, it may be quite useful in program evaluation. For example, it may indicate that, for whatever reason, no further effort in a particular area is necessary or that much remains to be done.

Whatever design is selected, it is important that information on program activities be related to estimates of program impact. For example, even in the absence of a group of students who have not been exposed in some manner to the program (i.e., no control group), it is generally possible to identify the level or "amount" of program exposure for each student or class, and treat those with the lowest exposure level as a comparison group relative to those with the highest exposure level. Comparisons between the gains made by the high group and those made by the low group would then provide the basic data needed for assessing whether or not the program had had any impact on the outcome measures, providing that there are meaningful differences between the activities engaged in by the two groups. If the activities are quantifiable on a fairly fine scale, such as number of hours, the "high" group might consist of the 25% of the students who had the greatest number of hours and the "low" group might consist of the 25% of the students who had the fewest hours of exposure. If the differences are in fact meaningful, the design would lend itself to analysis to determine the strength of the relationship between program activities and outcomes.

Analysis Alternatives

As indicated earlier, the selection of specific analytical techniques should not be separated from decisions regarding design and instrumentation. The appropriate technique is also related directly to sampling decisions; sometimes decisions regarding analytical techniques dictate sampling decisions and at other times they must flow from them.

Choosing the appropriate statistical technique for analysis requires specialized skill and knowledge. While responsibility in this area should be given to a specialist, the program manager should have at least a basic understanding of the more important factors in making the selection. This will facilitate dialogue between the specialist and the program staff, enhance the manager's ability to participate meaningfully in design, sampling, and instrumentation decisions, and be an aid in interpreting the results of the evaluation.

Three questions are central to the selection of appropriate analysis techniques. One concerns the level(s) of data being analyzed - - whether the data are nominal, ordinal, or interval in nature. Each of these levels of measurement describes various qualities connected with the data which have an effect on the type of statistical technique one can employ.

The second question concerns the number of variables being investigated. The type of statistics used to analyze only one variable differs from the type used to analyze two or more variables; that is, there are univariate, bivariate, and multivariate techniques.

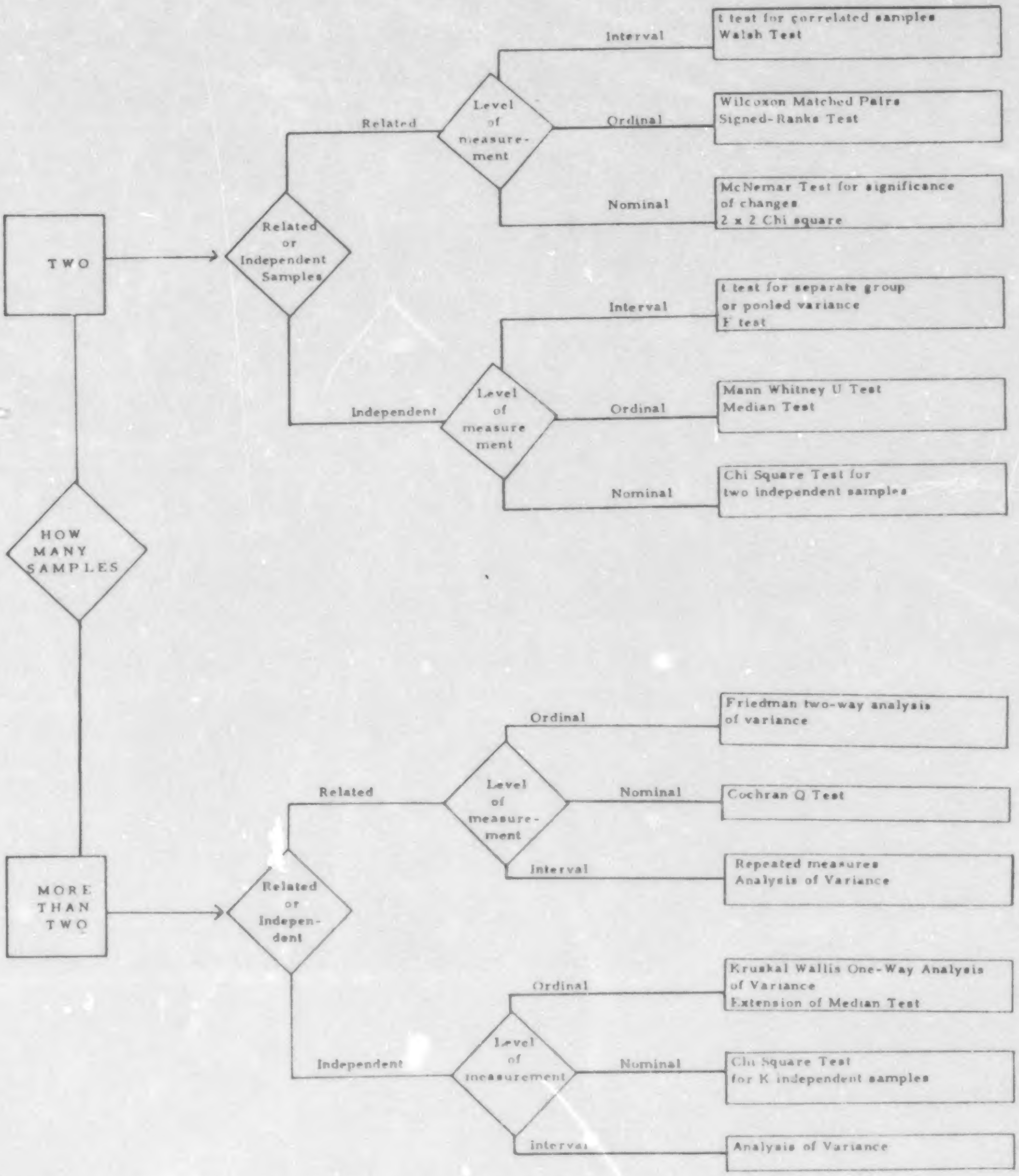
The third question is concerned with the nature of the evaluation question asked. Depending on whether you wish to compare the variables with each other, determine if a relationship exists between them, or merely describe the way the data are distributed, different statistical techniques will be applicable.

In summary, the selection of the appropriate statistical techniques requires answers to these three different questions. The three figures which follow illustrate the impact of these apparently simple questions on the selection of the appropriate analytical techniques. The techniques listed in the figures are not necessarily all-inclusive, but are appropriate given the conditions cited.

Each of the three figures starts from the question most appropriately answered by the program manager:

> "Is the interest in making comparisons, determining if there is a relationship, or describing the data?"

Figure 4 is applicable if interest centers on comparing two or more groups on some measure; for example, test scores of participating and non-participating students. The diamonds represent points at which certain decisions must be made. The first diamond asks, "How many samples are you dealing with?" If your answer is two, then use the upper portion. The next diamond, Related or Independent Samples, refers to the types of samples you are working with. Here we mean independence both within and between samples, for example, samples cannot be matched or paired as is the case in related samples. The third diamond refers to the level of measurement one is using: nominal, ordinal, or interval. If the data are ordinal, then one follows the line marked "ordinal" and at the end finds an appropriate statistical technique.



TWO
Related or Independent Samples
Related
Level of measurement
Interval
t test for correlated samples
Walsh Test
Ordinal
Wilcoxon Matched Pairs Signed-Ranks Test
Nominal
McNemar Test for significance of changes
2 x 2 Chi square
Independent
Level of measurement
Interval
t test for separate group or pooled variance
F test
Ordinal
Mann Whitney U Test
Median Test
Nominal
Chi Square Test for two independent samples
HOW MANY SAMPLES
MORE THAN TWO
Related or Independent
Related
Level of measurement
Ordinal
Friedman two-way analysis of variance
Nominal
Cochran Q Test
Interval
Repeated measures Analysis of Variance
Independent
Level of measurement
Ordinal
Kruskal Wallis One-Way Analysis of Variance
Extension of Median Test
Nominal
Chi Square Test for K independent samples
Interval
Analysis of Variance

FIGURE 4
COMPARE

Figure 5 is applicable if interest centers on determining whether or not two or more variables are related to each other in some way. For example, if there is interest in determining whether the number of hours of inservice training is related to teacher attitudes toward career education, the upper portion of the figure would apply.

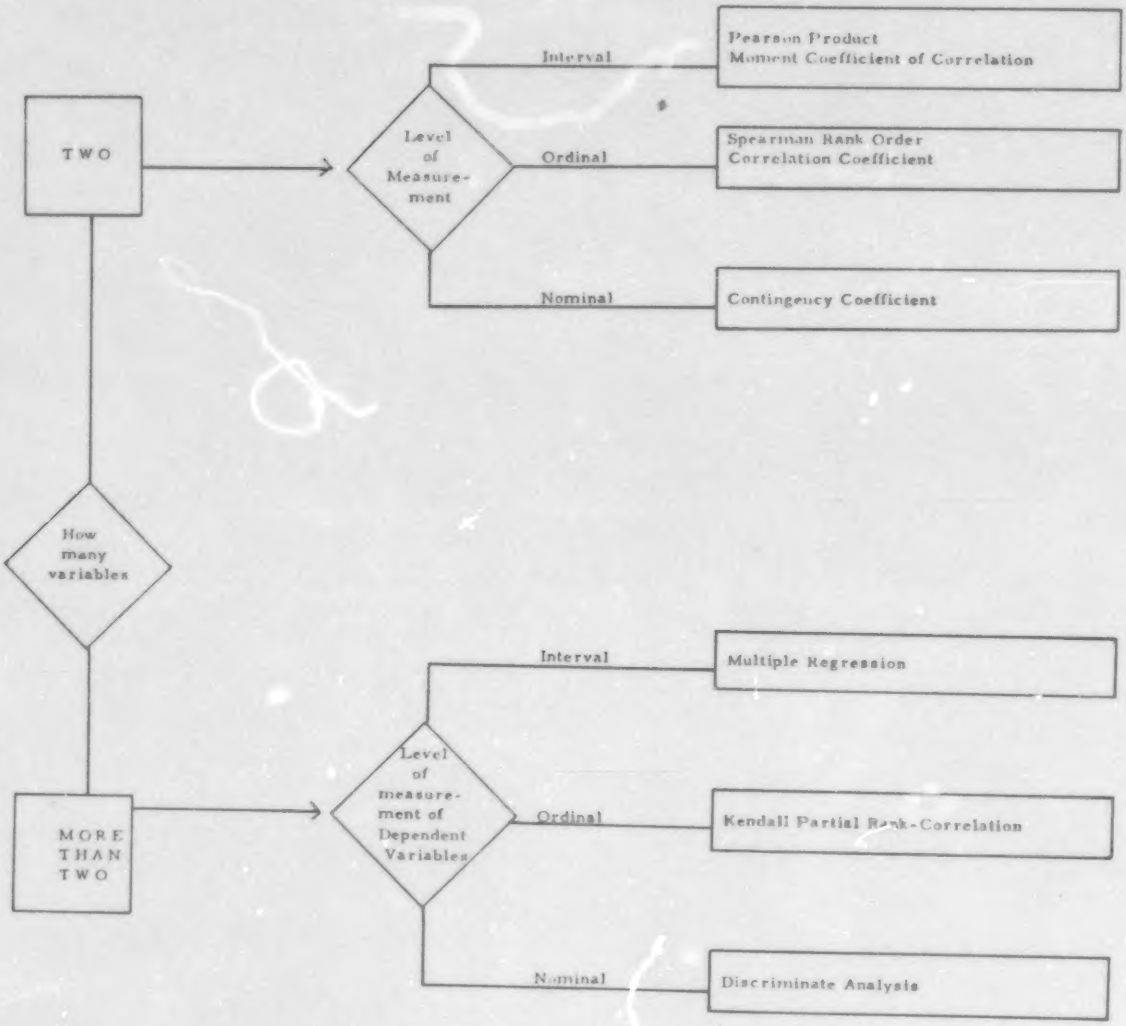

TWO
Level of Measurement
Interval
Pearson Product Moment Coefficient of Correlation
Ordinal
Spearman Rank Order Correlation Coefficient
Nominal
Contingency Coefficient
How many variables
MORE THAN TWO
Level of measurement of Dependent Variables
Interval
Multiple Regression
Ordinal
Kendall Partial Rank-Correlation
Nominal
Discriminate Analysis

FIGURE 5
RELATE

Figure 6 is applicable if interest is merely in describing what the data looks like. For example, if one had survey results showing the way different groups of teachers responded to questions about the quality of curriculum units (ordinal scale), then the median would be an appropriate statistic. This approach can be used with any set of data and hence begins with a question on the level of measurement.

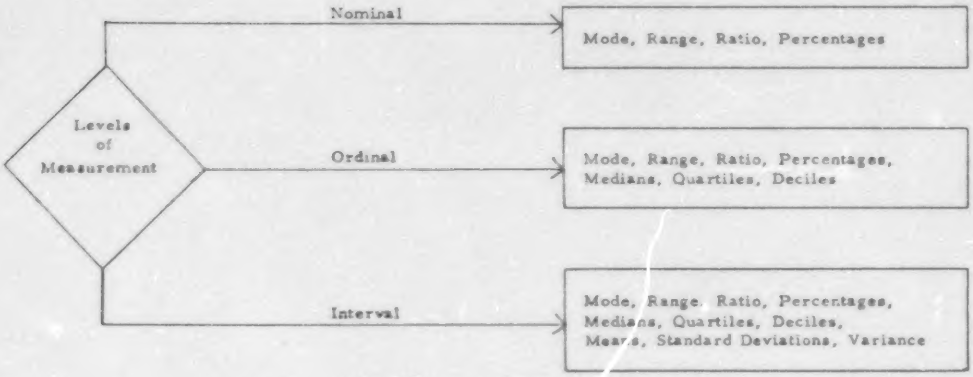

Nominal
Mode, Range, Ratio, Percentages
Levels of Measurement
Ordinal
Mode, Range, Ratio, Percentages, Medians, Quartiles, Deciles
Interval
Mode, Range, Ratio, Percentages, Medians, Quartiles, Deciles, Means, Standard Deviations, Variance

FIGURE 6
DESCRIBE

Summary

Early consideration should be given to the fixing of responsibility for design and analysis. This responsibility should normally be given to individuals with specialized technical training and experience. Provision must be made, however, for the participation of program staff to assure full consideration of program needs and constraints. Further, to provide for adequate balance between technical and program considerations, the program manager and his staff should participate in the decision-making process and should have an opportunity for review of significant decisions.

Decisions with respect to design and analysis should be tied tightly to the evaluation questions and to decisions regarding sampling and instrumentation. At this point in the evaluation process, relatively firm decisions should be made regarding design and analysis, but final decisions should await decisions in the area of sampling. It may be that constraints in the area of sampling may force some modifications.

CHAPTER NINE

PREPARE SAMPLING PLAN

Once a tentative decision is made on an evaluation design, the next step in the process is to prepare the sampling plan. The plan should identify the individuals from whom information will be collected, how and when they will be selected, and approximately when data collection will be carried out (e.g., "mid-September" and "mid-May").

Sampling is a procedure in which a part (individual students, classrooms, schools, etc.) is selected from the whole (the school, school district, etc.) in such a way that it mirrors the population and allows generalizing the results of the sample to the population. Of primary concern is determining how the sample should be selected since the sampling approach used relates directly to the power of the generalizations which can be made as well as to the design and analysis plan.

One of the initial concerns should be the fixing of responsibility for the sampling plan. Usually this responsibility is assigned to the person(s) responsible for design and analysis because of the close interrelationships among the selection of the sample, the development of the design, and the analysis of the findings. The individual carrying out these responsibilities should possess statistical expertise because of the sophisticated techniques involved in sampling. If no member of the program staff has the necessary training and skills, the services of an outside statistician should be obtained. In preparing the plan there will be need for significant inputs from the program manager and his staff to assure that the plan is consistent with local realities. This should include participation in the process of determining the sample and review of the final plan.

In an educational setting there are often many restrictions on the selection of samples; this in turn has important implications for the design and analysis approaches which can be used. Indeed, in some cases, issues of sampling may lead to reconsideration of design alternatives.

One area of potential difficulty relates to the "randomness" of the sample. Educational research usually seeks the random assignment of students to program activities (treatments) prior to program exposure, and designs which do not involve the assignment of students to program activities by means of random or some other form of probability sampling procedures will be suspect. It is recognized, however, that local conditions may prevent such assignment, and in such cases, a commonly followed approach is to randomly select a sample from among existing groups (e.g., participating and non-participating students). Where this procedure must be followed, it is important to investigate and record in as much detail as possible the manner in which the groups were formed to assure the optimum representativeness of the sample (i.e., were certain types of students encouraged to enroll in program classes? What made teachers volunteer for program training sessions? etc.). The results of this investigation are essential to judging the extent to which findings can be atributed to program activities rather than specific characteristics inherent in the sampling groups.

Another problem relates to the size of the sample. In many program evaluations it is difficult to select a large enough number of students from a school or from preselected classrooms within a school and still assure representative selection. In this situation, often it is necessary to randomly select schools or intact classrooms within schools. If it is decided to randomly select classrooms rather than students, then the unit of evaluation becomes the classroom, rather than the student. If four classrooms of 25 students each are selected, the size of the sample will be four, rather than 100. A more complete discussion of the appropriate experimental unit in statistical analysis can be found in numerous texts on experimental design and statistical analysis. The point here is to indicate that the unit of analysis will vary depending on selection considerations.

In general, sampling requires a knowledge and implementation of procedures which allow for an unbiased selection of a small group that is characteristic of the larger population such that estimates based on this smaller group can be generalized to the larger group from which it has been drawn. Assuming that an adequate sample can be drawn, considering the design and unit of analysis, the next and most frequently asked question is: "How big a sample must I take?" Since accuracy and power are increased by sample size, a larger sample would be better than a smaller one. In other words, with all other factors held constant, the larger the sample, the smaller the sampling error. The underlying question is, "How much of a sampling error are we willing to put up with?" The answer depends upon: (a) the consequences of reaching wrong conclusions; and (b) the cost of reducing the sampling error.

The decisions involved in answering the above questions cannot be made on the basis of statistical or methodological considerations alone, since assessment of the consequences of errors of given magnitudes and cost factors is essentially an administrative consideration. Further, they depend upon the distribution of the variables in the population, and this is generally not known. As a reasonable approach to these problems, most program evaluations develop sampling plans which call for samples of approximately 5% (with a minimum of 30 if the group is relatively small) where the population numbers up to 5,000. For larger populations, i.e., over 5,000, a sample size of approximately 2% is generally considered adequate. In other words, sample sizes of between 30 and 250 where the population ranges up to 5,000 and of at least 100 for populations in excess of 5,000 are generally viewed as being stable enough to insure a sufficiently low error of estimate with a reasonable expenditure of time and money. Of course, these sample sizes may vary in accordance with the specifics of the evaluation design. Either the 90% or the 95% confidence level is utilized to judge the significance of the findings.

The preceding rules of thumb with regard to sample size apply to each "homogeneous" group, e.g., students in the 9th grade who comprise participant group X, or students in the 9th grade who constitute a "no treatment" group. They also refer to samples in which the students in the sample have been randomly selected. In relatively small districts where intact classrooms have been selected, i.e., where classrooms rather than students constitute the unit of analysis, it is best to select as many classrooms as possible (i.e., at least 10 to 20) in order to maximize the stability of the findings and to maximize the sensitivity of the statistical techniques which are applicable to small samples.

In summation, program personnel considering sampling issues should keep the following in mind:

- Sampling is not an issue when all, i.e., a census, of the population is tested.
- Generally, the larger the sample, the better the measure.
- Typically, random procedures should be followed in selecting a sample.
- Confidence levels and therefore sample sizes can vary based upon practical or administrative considerations.
- As a rule, for populations of 5,000 or under a sample size of 5% or 30, whichever is the larger, is considered an appropriate or safe sample.
- The 30 or 5% rule applies for each participant group.

- If classrooms are selected as the sample unit, it is likely that more students will have to be tested to insure a valid sample than if the student is the sample unit.

- In general, the design of the sample is a very sophisticated problem requiring the expertise of a statistician. If no member of your staff has sufficient background in this area, the services of an outside statistician should be obtained as an advisor.

CHAPTER TEN

FINALIZE DESIGN AND ANALYSIS AND PREPARE A REPORTING PLAN

Once the evaluation questions have been selected, instruments chosen or developed, and tentative design, analysis, and sampling plans made, it is necessary to review the evaluation design in its entirety in terms of costs and technical limitations. Once this is complete and any modifications made as a result of the review, the final step in the design process is to prepare a plan for reporting the results of the evaluation.

Design Review

Throughout the process of specifying questions and instruments, and settling on appropriate sampling and analysis strategies, a series of decisions have been made which possibly: (a) redefined somewhat the original conception of the final products of the evaluation; (b) affected the anticipated cost of conducting the evaluation. Before finalizing the design, it is necessary to review this series of decisions to insure that the results are acceptable. Systematic recording of basic decisions on the Evaluation Design Worksheet presented at the end of Chapter Two will greatly facilitate this process.

In most situations an important consideration in the design review will be the overall scope and cost of the effort. Among the major cost elements may be:

- purchase of tests;
- development of tests, questionnaires, and other instruments;
- scoring of tests (sometimes this must be done by the publisher or some other service agency; other times it may be done locally and in large studies might include purchase of computer time);
- staff costs of personnel added for the preparation, processing, and analysis of the study.

While in theory it might be preferable to decide what needs to be done and then to allocate the necessary resources to accomplish the task, in reality this is rarely, if ever, possible. Typically, a ceiling cost of the evaluation was determined before the details of the design were specified and the issue at this point is to ascertain whether the ceiling will or will not be exceeded. If projected costs are below the ceiling, then any decisions which limited the evaluation design because of cost should be reexamined. The objective should be to determine whether any improvements can be made by increasing the cost, involving reconstructing a larger sample, budgeting additional staff, or selecting more sophisticated testing instruments. If no further improvements can be made, the unneeded funds should be made available for other purposes.

If, on the other hand, the ceiling will be exceeded, then it may be necessary to compromise on some of the substantive products of the evaluation. This may require eliminating certain data categories because of the cost to obtain the information (i.e., community surveys, followup on students, etc.). Once the scope of the study and costs have been determined and any necessary changes or modifications of the design completed, it is possible to move to the next step of the review.

To assure that the study is responsive to program needs, the general categories and questions which the various interested parties in the career education effort will find useful in both planning and operating the project should now be examined. This review should focus on whether the design, as constituted, will provide the information necessary to answer the questions raised by the various parties concerned with the career education project, identified in Step 3. For example, some questions will yield rather direct and complete answers in certain categories, while other responses will be more indirect or inferential. The extent to which the design will satisfy the information needs or desires of all interested parties should be made explicit at this point so that there is complete understanding of exactly what kind of information will be obtained by the evaluation as well as how completely the questions can be answered.

When completed, this review procedure provides the basis for preparing the reporting plan.

Reporting Plan

The data reporting plan will depend on those who will be using the evaluation report. It is often best to assume that a wide range of persons will want to read the report. Besides local project personnel, the range of other interests might include:

- the U.S. Office of Education;
- the state education agency;
- the local board of education;
- school administrators at all levels;
- teachers;
- guidance personnel;
- parents and other community residents; and
- other career education project managers.

The report should be organized so that it can be easily interpreted by everyone who will want access to it. A sample format is provided here.

I. OVERVIEW OF CAREER EDUCATION EFFORT

The basis for the overview should be the document developed in Step 1.

II. EVALUATION QUESTIONS AND METHODS

In this section of the report the evaluation questions that the study addressed should be stated with a discussion of any problems or reformulation of the questions due to absence of data or procedural constraints.

This should be followed by a brief summary of the methods employed to obtain the results.

III. FINDINGS AND DISCUSSION BY QUESTIONS

This part of the report should list each evaluation question followed by two subsections:

1. Findings; and

2. Discussion.

(1) The findings are factual statements relating to the evaluation question and the information and answers resulting from the effort should be included. (2) The discussion subsection should contain an explanation and analysis of the findings. This would include, for example, the extent to which the question was capable of being answered and any elaboration on why certain information was not available. All charts or graphs indicating student outcomes and other project results referring to the question should be included.

IV. CONCLUSIONS

This part of the report should summarize the findings and analysis of the project as a whole. If any problem areas or deficiencies were identified in evaluating the discussed questions, they should be noted with suggested actions or recommendations for improvement.

APPENDIX

The appendix should contain copies of all instruments, questionnaires, etc., used for the evaluation.

In developing the outline for the report, attention should be given to actually constructing the tables that will be used. After these have been constructed, a review of the data tabulation forms should be made to assure that the information will be available for tabulation in the easiest possible manner. By examining these tables at this point, time can be saved in tabulation later on.

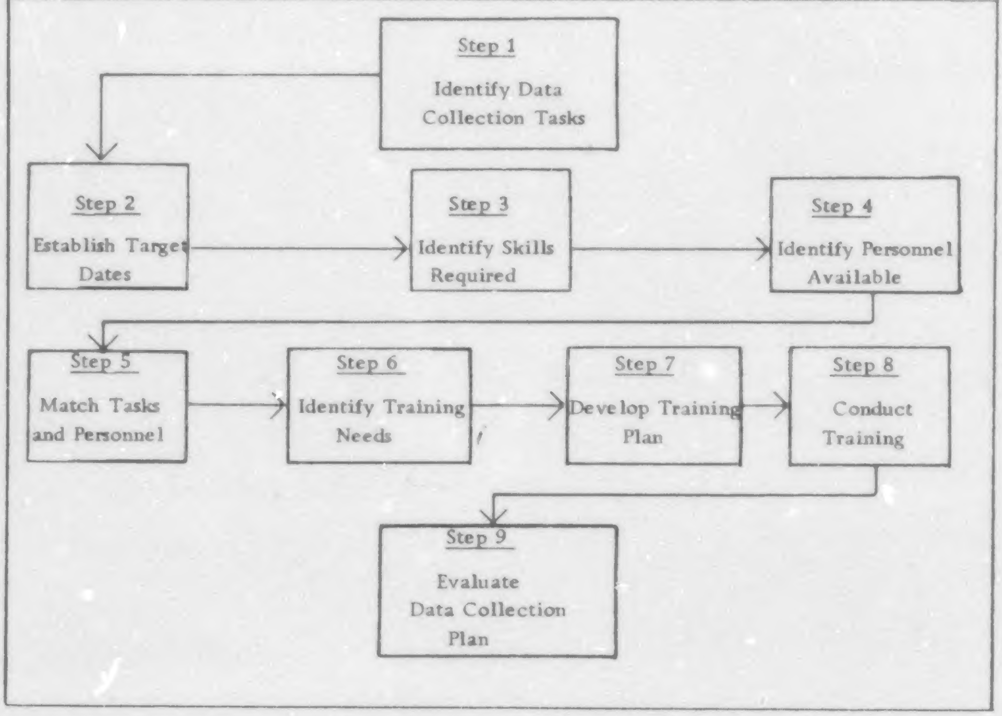
Depending on the evaluation design, it may be appropriate to provide interim reports. For instance, if information is being collected from teachers on a weekly basis regarding what they are doing in their classrooms, it may be helpful to report the findings of this activity at various intervals. There are two reasons for this. The information may be helpful to decision makers, and the teacher will respond better if he/she receives some feedback from her/his efforts. This is but one example of situations in which interim reports may be appropriate.

In addition to the comprehensive report, it may be desirable to plan on the preparation of special summaries for persons with particular areas of interest and whose interests in methodology are minimal.

CHAPTER ELEVEN
PREPARE DATA COLLECTION PLAN

The key to successful data collection is a well thought out plan that identifies the logistical activities in advance, provides a basis for logically selecting the personnel required to collect the data, and specifies time guides for judging progress. Table 11.1 illustrates a nine-step plan which can aid in designing a data collection effort appropriate for your career education evaluation study. Tables 11.2, 11.3, and 11.4 are provided at the end of this chapter as aids in developing the plan. Below is a graphic illustration of the overall plan.

TABLE 11. 1


Step 1
Identify Data Collection Tasks
Step 2
Establish Target Dates
Step 3
Identify Skills Required
Step 4
Identify Personnel Available
Step 5
Match Tasks and Personnel
Step 6
Identify Training Needs
Step 7
Develop Training Plan
Step 8
Conduct Training
Step 9
Evaluate Data Collection Plan

The first step in designing a data collection plan is to develop a list of all the tasks involved (the tasks may be listed in column 1 of Table 11.2). In preparing the list, think of a task as anything that someone will have to do to carry out the data collection process. Do not overlook simple activities such as the reproduction of materials. Often these small tasks cause major problems for the entire evaluation effort. In recording all the tasks, list them in sequential order so that nothing will be overlooked. When finished, carefully examine the list to see that nothing is missing and that tasks flow logically from one to the other. Be sure that you are not expecting some activity to occur without having a task somewhere along the line to carry out that activity.

The second step is to establish starting and completion or target dates for each of the tasks. (Starting dates, which may be entered in column 2 of Table 11.2, will be the same for many activities; some will come later and be dependent on completion of other activities. The completion dates may be entered in column 3 of Table 11.2.) These dates should be realistic and, if possible, provide some flexibility or cushion to allow for unexpected delays. Be sure to take into account holidays, vacations, school activities, etc., when assigning completion dates to the tasks. With step 2 completed, you should know what has to be done and when.

Step three involves the identification of the skills needed to complete each of the tasks. (These may be recorded in column 2 of Table 11.3.) One example of such a skill might be experience in test administration. Another might be experience in telephone interviewing.

The fourth step involves identification of the personnel available to carry out the data collection. (These may be recorded in column 1 of Table 11.4.) You will need to know the constraints on the use of various personnel. Some will have only certain hours or days available to work on the data collection. Another consideration will be the cost. Some personnel will have to be paid out of project funds. In these cases, the budget will have to be considered in deciding who to use. In addition, an assessment of the skills possessed by the available personnel should be made. (Constraints on the use of personnel may be recorded in column 2 of Table 11.4 and notation of their relevant skills in column 3.) In larger jurisdictions where it may not be feasible to handle all of this detail centrally, it may be desirable for planning purposes to provide for summarization of some of the information in Tables 11.3 and 11.4.

The fifth step involves matching the personnel and tasks using the information you have developed in the previous steps. In some cases you may want to assign prime responsibility to one person, and also additional personnel for support activities. (These assignments may be recorded in columns 4 and 5 of Table 11.2.)

The sixth step is to determine the training needs of the personnel you have decided to use. This can be done simply by examining the required skills as you have outlined them in step three (Table 11.3, column 2), and matching them to the skills of the personnel you have selected. Discrepancies between the two will have to be resolved by training the personnel to meet the project needs. (The discrepancies may be recorded in column 3, Table 11.3.)

The seventh step requires the development of a training plan for providing the identified skills in the previous step.

The eighth step is to implement the training plan. In cases where you are using instruments with which personnel are unfamiliar, there should be a training session to familiarize the staff with the instrument. An example of an outline for such a training session is:

A. Distribute samples of the test and answer sheets to the trainees.

B. Discuss the test, explaining all relevant aspects.

C. Give the trainees copies of blank answer sheets or use a device for projecting the information for all to see. Discuss the necessary identification items on the answer sheet and the methods by which the item responses are to be indicated.

D. Administer all or part of the test to the trainees and when feasible have trainees administer parts of the tests (for critique by instructor and other trainees).

E. Discuss the results and the interpretation of the instruments.

F. Clarify any questions the trainees have relative to the instruments.

G. Discuss the advantages and disadvantages of this particular instrument and any particular problem the trainee might have in administering the instrument.

Having completed the above eight steps, you have almost completed your data collection plan. The plan should be thought of as flexible and as a guide. When changes must occur, do not abandon the entire plan, but change it to meet the new needs. A good data collection plan can be a real time saver but it almost always goes through a number of revisions.

In some instances all or part of the data collection will be completed by an outside evaluator. When this occurs, steps 4 through 8 will be discussed by the evaluator in the proposal submitted to the program. In reviewing such proposals the program staff should note the degree to which the points raised in these steps are realistic.

When outside evaluators are utilized, it will be necessary to determine the cost factors involved in carrying out the plan. If no contractor is willing to bid on the proposal, then either the scope of work must be altered or the budget increased or efforts expanded to recruit interested evaluators.

As a last step you should evaluate the data collection plan after data collection has been completed and note any shortcomings that should be avoided during subsequent evaluation efforts.

TABLE 11. 2
DATA COLLECTION PLANNING CHART

(1) TASK	(2) Starting Date	(3) Completion Date	(4) Prime Responsibility	(5) Support Responsibility

TABLE 11. 3
TRAINING NEEDS ASSESSMENT

(1) TASK	(2) Skills Required	(3) Skills Needed by Staff Members	
		Skill	Staff

TABLE 11. 4
PERSONNEL PLANNING FORM

(1) Name	(2) Restrictions - Cost, Availability, etc.	(3) Skills

CHAPTER TWELVE
IMPLEMENT THE EVALUATION

Introduction

Despite all the careful advance preparation, implementation will not "just happen." It will require assignment of clear-cut responsibility for the management of implementing the evaluation, including data collection, analysis, and reporting. The need for attention to the management of implementation is important because of the numbers of people who will play some part in data collection and the variety of things that can go wrong over a period of a school year. It is important also because of the time-phasing of the evaluation process which will normally occur, with peak activity near the beginning and end of the school year with a routine flow of information in between. This uneven and irregular type of activity requires careful monitoring.

Management of the implementation process will include such matters as:

- assuring wide distribution of the program manager's authorization of the evaluation study;
- advance checking to make sure individuals are prepared to carry out their assigned responsibilities;
- following up on performance to make sure work was on schedule and properly performed;
- when data reflect misinterpretations, making sure corrections are made promptly and instructions are clearly understood for future submittals;
- watching for bottlenecks and establishing a smooth flow of work;
- clarifying misunderstandings of plans and instructions; and
- handling crises.

Data Collection

Application of the data collection plan outlined earlier is the key to sound data collection efforts. The planning chart developed as a part of that plan provides a task-by-task list of each activity to be carried out. It also provides you with starting and target or completion dates, and names the staff members who have prime responsibility and support responsibility for each of the tasks. The data collection plan also includes a training needs assessment by named individuals and a training plan for the staff involved in the data collection. The training should be implemented at least a week before the first data collection takes place and alternative arrangements should be made for training of those who miss the regular training session. In some cases arrangements may have to be made to provide separate training for persons with different evaluation responsibilities, who work a different schedule, or are physically separated from the rest of the evaluation personnel.

Basically, there are two types of data to be collected these relate to:

activities that occur only one time during the course of the evaluation such as pretest, posttest, certain observations and interviews; and

activities that are continuous and usually spread over the life of the project. An example of one such activity is the regular reports that a teacher might submit as a part of the evaluation. These reports are repeated over and over on a routine basis.

It is recommended that the evaluation staff routinely monitor the data that is being submitted. There are two reasons for monitoring this information. The first is to assure the quality of the data being collected. If the quality goes down, corrective measures can be taken that will provide a higher quality of data. The second reason is to give feedback to the person collecting and submitting the data. The feedback is often helpful to the person providing the data and makes the person more likely to cooperate in all aspects of the evaluation because he or she knows that someone is interested enough to look at the reports and check back for clarification or correction.

Hopefully, all the tasks related to the collection of data will be completed on time and without any difficulty. However, in an evaluation process, various events can occur that are outside the influence of the evaluators. The evaluators must always try to anticipate these events and to be flexible enough to overcome these obstacles. The following are some examples of typical problems that may occur:

- a flu epidemic the day of the tests;
- a lack of cooperation from participating staff or students;
- personnel who have been closely involved in the evaluation are suddenly transferred;
- new program activities are added that impact on certain outcomes; and
- budgetary changes are necessary that will reduce evaluation monies.

Each of these kinds of problems can have a serious effect on the evaluation study. Of course, there are only a certain number of precautions that one can take. For example, by having staff closely involved in the design of the evaluation, the chance that staff will not cooperate is minimized. One cannot, however, predict the flu. Handling these problems requires the best judgment of the staff. In some cases, one might have to modify all or part of the design. In making any change in the evaluation design, it is important to record accurately exactly what did occur so that there is a record to refer to during the analysis phase. This record of deviations from the original design should be included in the evaluation report so that those using the information contained in the report understand the context in which the findings are to be considered.

The last step in implementing the data collection plan should be an evaluation of the plan and all data collection activities. Often, people involved in evaluation make the same errors year after year. Ideally, every person involved in the data collection should be involved in the evaluation. If a number of the "bugs" in the data collection plan become apparent on the day that the major collection activity is taking place, these should be noted promptly. Other "bugs" will turn up when the data are tabulated. An evaluation will note any deficiencies and pinpoint or suggest their causes.

Data Processing and Analysis

The processing needed to prepare the data for statistical analysis should be implemented as soon as possible after the first batch of data has been collected. This would include any necessary labeling for identification purposes, coding, and key-punching, if machine processing is to be used. The process should be undertaken early for several reasons: a) the system governing the flow of the data can be improved early; b) certain defects in the data can be corrected while the facts and circumstances are fresh in people's minds; and c) labeling or coding difficulties can be discovered early, minimizing the need to make corrections on large amounts of data which have already been processed. Insuring that labeling is error-free is of particular importance in a pre/post type of study because of the need to link or pair data from the posttest phase with data from the pretest phase on a student-by-student and a class-by-class basis. This aspect of data processing should be carefully developed or reviewed by the computer programmer if machine analysis is to be used, or by the statistical consultant if the analysis will be done manually.

It should be anticipated that the realities of the data collection effort are likely to require modifications in the data analysis plans; these would stem chiefly from deviations from the original sampling plan and from the fact that incomplete data will have been obtained from various students or various classes. Techniques exist for handling or accommodating most of these deviations, and it is essential that the analysis specialist be kept informed of such deviations as they occur. Modifications in the data analysis plan can also be anticipated on the basis of preliminary tabulation of early data; such tabulation would display basic facts about each measured variable, such as means, standard deviations, ranges, and sample sizes for each homogeneous grade and participant group. Other descriptive statistics may be useful for internal program use and/or for refining the data analysis plan. The procedures for obtaining these and all other tabulations would be developed by the evaluator or the statistical consultant.

It should be pointed out that most studies in education leave insufficient time for carrying out the data analysis. The more time that can be allocated for analysis, the greater will be the opportunity to pursue any leads, hunches, or questions which each phase of the analysis might suggest. The plan, in other words, should be viewed as providing the basic framework for analysis, rather than as a rigid prescription. In this way, the data will yield the maximum degree and type of meaningful information and important clues for interpretation or explanation of the basic findings.

Reporting Findings

In the broadest sense, the purpose of evaluation is to provide information concerning performance to decision makers and planners preparing for the future. The evaluation process is not complete until the knowledge gained is available and utilized by planners, managers, and policy makers. This, of course, is as true for career education programs as it is for others. Because career education is a relatively new effort, and a large body of program information has not yet been developed, the reporting and dissemination of evaluative information is especially important for individuals seeking to improve career education programming at all levels.

Different kinds of information are needed at different levels of planning and administration, and the reporting plan should take this into account. As a rule, individuals at the operating level (teachers and counselors) require specific information on each student (individual data) while policy makers at the school board level need more general information (aggregate data). At the state and federal level the information need becomes increasingly general. Clearly, policy makers and planners at higher levels do not need, nor could they handle, detailed and specific data from many sources. The need for increasingly general information at higher policy making levels carries with it a need to compare programs in terms of their settings (context), objectives, activities, participants, and results. The evaluation and reporting plans should be designed to provide information which will be useful at various levels of the career education program.

Where possible, local objectives and activities should be related to broadly used categories (e.g., those outcomes cited in the tables in the Appendix, standard categories used by the state, etc.). This makes it more possible for planners and policy makers to compare types of programs and results without being faced with the impossible task of trying to make judgments based upon many individual case histories.

To assist those persons not familiar with the context in which the career education effort was evaluated, the inclusion in the report of certain general, descriptive information would be helpful. Particularly valuable in this regard would be information regarding the size of the school district, and the socioeconomic characteristics of the area.

In state or federally funded programs, the funding source will usually stipulate the kinds of information to be reported. In general, however, it is important to describe changes that have occurred in the target populations. It is also important to describe the relationships between planned change and those changes which actually occurred. In reporting change, it is necessary to describe what produced that change. In other words, what was different for one student or group of students than for others; or, what was different this year as compared to last year? Without this descriptive information it will not be possible to understand the significance of any changes that do occur. Information that relates to the difference between planned and actual change is important in understanding which activities or approaches are most effective, etc. The value of this kind of information to planners at all levels should be clear.

Writing a report without disseminating that report is wasteful. One should consider disseminating results to as many groups as possible. Even if your program is not funded by USOE, it would be worthwhile to send them a copy. Finally, you are strongly urged to send a copy of your final report to the ERIC Clearinghouse for Career Education (Northern Illinois University) so that it will be available to interested persons throughout the country.

APPENDIX

SUMMARY OF RESULTS

CAREER EDUCATION INSTRUMENT REVIEW
CONDUCTED FOR
U.S. OFFICE OF EDUCATION

TABLE OF CONTENTS

Page

I. INTRODUCTION a-3

II. PANEL PROCEDURES a-4

III. RESULTS OF THE REVIEW a-6

A. Findings a-6

Table I: Recommended Instruments and Instruments Considered "Promising" a-7

B. Information Concerning Individual Tests a-10

IV. CAREER EDUCATION STUDENT OBJECTIVES a-27

Table II: General Career Education Learner Outcomes: USOE Policy Statement a-28

Table III: General Student Outcome Areas and Specific Objectives Applicable for Round II Projects Funded Under Part D of Vocational Education Act a-29

Table IV: Student Outcome Objectives of the Experience-Based Career Education Programs (EBCE) a-31

V. LIST OF PANEL MEMBERS a-35

I. INTRODUCTION

In 1974 and 1975, an extensive nationwide search for instruments which measured career education objectives was conducted by Development Associates in an attempt to provide an instrument review panel with as many tests as possible for review. In selecting instruments for review, the following criteria were applied:

> The test could be administered on a group basis.
>
> Student responses required little or no writing.
>
> The instrument's objectives could be related to one or more of the objectives of career education (see Tables II-IV, Section IV).
>
> Each instrument was available to local school districts for use in program evaluation.

One hundred and thirteen tests were reviewed against these criteria by panels meeting in August 1974 and August 1975. As a result of this review, ten tests were recommended and four were considered "promising."

II. PANEL PROCEDURES

Two panel meetings were held in the Washington office of Development Associates. The first, convened in August 1974, included the following members:*

Dr. Nancy Burton, Education Commission of the States;

Dr. William E. Coffman, University of Iowa;

Dr. Edward Lareau, Admiral Peary Area Vocational-Technical School;

Dr. Dale Prediger, American College Testing Program; and

Dr. Donald Super, Columbia University.

This panel reviewed all available materials and recommended several instruments for use in programs of career education.

In August 1975, a second panel met for two days at Development Associates' Washington office to review 23 additional tests that had been collected. This panel consisted of the following members:*

Dr. William E. Coffman, University of Iowa;

Dr. John O. Crites, University of Maryland;

Dr. Lois-ellin Datta, National Institute of Education;

Dr. Donald Super, Columbia University; and

Dr. Bert Westbrook, North Carolina State University.

Panel members reviewed instruments individually and after presenting the results of their reviews, the panel reached consensus on a given instrument. Recommendations were given as to whether or not the instrument should be recommended for use, recommended as "promising," or not recommended for inclusion in this guide for the evaluation of career education.

---

*See Part V for professional background information on panel members.

In addition to the review by the panels in September 1975, each of the recommended tests and the tests considered "promising" was reviewed by two project directors: Ms Ellen Poole, Supervisor of Career Education, Petersburg, Virginia, and Dr. Edward Lareau, Director of the Career Education Project, Ebensburg, Pennsylvania. They found each of the instruments satisfactory from the point of view of relevance and administrative feasibility.

If a test is "recommended" for use, it is considered by the panel to be a valid and reliable instrument. That is, the test contains good psychometric data, is easily administered and scored, and will measure clearly certain career education objectives.

If a test is listed as "promising" in the view of the panel, it needs further psychometric development (i.e., reliability, validity, and norm data). In addition, there may be some problems in scoring, or stereotyping on individual items. With these limitations in mind, it is suggested that these tests be used over those without any psychometric data or those "home-made" instruments lacking proper development.

Certain tests are not included in this guide because the panel considered them either lacking appropriate psychometric data, or inappropriate for the purpose of measuring career education objectives.

III. RESULTS OF THE REVIEW

A. Findings

Table I summarizes the results of the panel review. It presents the names of all recommended tests, what the panel judged they measure, and how they correspond to objectives in USOE's Career Education Policy Statement, earlier identified career education objectives, and objectives of experience-based career education models. In addition, selected comments of the panel concerning the tests are noted. It also contains a list of instruments which were considered by the panel as "promising."

Table I can be used in the following way. The column entitled "Test Outcome Objectives" tells what the panel judged the test measures. For example, the Career Planning Test of the New Mexico Series (cited as promising) measures school and career problem solving. The next three columns provide the user with additional information. That is, this test will also aid in measuring objective 6 of Table II, objective 5-e of Table III, and objectives IV-1, 2, 3 of Table IV. Thus the numbers and letters found in these columns are specific references to the objectives listed in Part IV of this summary. The column heading tells you what table in Part IV to use and the numbers and letters refer you to specific objectives.

Following Table I, additional information is provided on each of the tests listed. This information is given in order to provide possible users with more detailed data with respect to ordering and using these tests.

TABLE I
TEST SUMMARY SHEET FOR
RECOMMENDED INSTRUMENTS

Test	Test Outcome Objectives	Grade Level	General Career Education Learner Outcomes: USOE Policy Statement [a]	Student Outcomes Applicable for Round II, Part D VEA ([b])	Student Outcomes for Experience-Based Career Education ([c])	Remarks
Assessment of Career Development:		8-11				• Should not be used below 8th grade except for groups having above average verbal ability; • Answer sheets provide for collecting responses for up to 19 locally constructed questions.
Subscore I	Occupational Characteristics		--	4-a, b	I-4	
Subscore II	Occupational Preparation Requirements		--	4-c	I-4	
Subscore III	Exploratory Occupational Experiences		--	4-a 5-b	I-4 II-1	
Subscore IV	Career Planning Knowledge		4	4-d 5-c, e	I-2	
Subscore V	Career Planning Involvement		4 4	5-c, f 7-a	I-2; I-8; I-2; II-1	
Career Development Inventory:		8-12				There is sex stereotyping in a few items, but empirical studies show no sex bias; revision to remove stereotyping is in process.
Scale A	Planning Orientation		4	5-e, f	I-8	
Scale B	Resources for Exploration		4, 5	5-c, f 7-a	I-8 I-2; II-2	
Scale C	Information and Decision Making		4	1-d; 4-a, c; 5-c, e	I-4 I-2	
Career Maturity Inventory:		7-12*				• Minor degree of sex stereotyping; no evidence of sex bias; • The Attitude Scale items may permit the development of meaningful subscores.
Attitude Scale	Attitude Toward Work		3	3-b	I-3	
Competence Test:						
Part 1	Self Appraisal		5	4-a, e 5-a, b, e	I-4; I-5; I-1; II-1	
Part 2	Occupational Information		--	4-a	I-4	
Part 3	Goal Selection		5	5-a 5-b	I-1; II-1	
Part 4	Planning		--	4-c	I-4	
Coopersmith Self Esteem Inventory (long form)	Self Esteem	Age 9-adult	--	1-c	--	
Differential Aptitude Test		8-12	--	5-a, b	I-1; II-1	• Career Planning Program is currently formulated on an individual student basis; publisher is prepared to offer a group summary of findings • Might be used in lieu of achievement testing for outcomes 2a, b, Table III.

---

[a] Table II, Part IV.

[b] Table III, Part IV.

[c] Table IV, Part V.

* Sixth grade reading level.

TABLE I
TEST SUMMARY SHEET FOR
RECOMMENDED INSTRUMENTS -- cont.

Test	Test Outcome Objectives	Grade Level	General Career Education Learner Outcomes: USOE Policy Statement a/	Student Outcomes Applicable for Round II, Part D VEA (b/)	Student Outcomes for Experience-Based Career Education (c/)	Remarks
How I See Myself	Self Esteem	3-12	--	1-c	--	• May encounter some scoring difficulties; • Reliability of factor structures at the elementary level is somewhat questionable.
New Mexico Career Education Test Series: Career Oriented Activities	Exploratory Activities	9-12	4	5-c, f	1-2	
Piers-Harris Children's Self Concept Scale	Self Esteem	3-12	--	1-c	--	Further data needs to be collected on the subscores suggested by the factor analysis.
Self Observation Scale Subscales: K-4 5-12	Self Esteem	K-4 5-12	--	1-c	--	• This test was recommended by the review panel only on the condition that the publisher provide users with information regarding the scoring keys and weights for the subscales. • Local review for sensitive items recommended by review panel. • Further work on validation needed.
Youth Inventory	Self Esteem; Career Awareness; Career Attitudes	7-12	2, 6, 9	1-c; 5-b; 6-a; 7-a	1-6	
TEST SUMMARY FOR INSTRUMENTS CONSIDERED PROMISING						
Employment Readiness Scale	Readiness for Employment; Work Values	11 & 12-adult	3	3-b	--	• Test is especially applicable for high school students who will begin working upon graduation; • Test needs some further psychometric development; • Low correlations among test items.

---

a/ Table II, Part IV.

b/ Table III, Part IV.

c/ Table IV, Part V.

TABLE I
TEST SUMMARY SHEET FOR
INSTRUMENTS CONSIDERED "PROMISING" -- cont.

Test	Test Outcome Objectives	Grade Level	General Career Education Learner Outcomes: USOE Policy Statement [a]	Student Outcomes Applicable for Round II, Part D VEA ([b])	Student Outcomes for Experience-Based Career Education ([c])	Remarks
Minnesota Cognitive Questionnaire for Career Education	Occupational Information	1-3 4-6 7-9	--	4-a, c	I-4	• Some sex and ethnic group stereotyping on written test items and illustrations; • Incorrect factual content on some items; • Forms overlap so that discrimination ability is lessened between grade levels; • Test needs to be scaled on a longitudinal basis.
New Mexico Career Education Test Series:		9-12				
Career Planning	School and Career Problem Solving		6	5-e	IV-1, 2, 3	
Job Application	Ability to Apply for Jobs		4	7-b	I-7	
Career Development	On the Job Skills Adjustment to Work		2	6-b, e, f	I-5 VII-1, 2	
Orientation of Career Concepts Tests 1-3 Tests 5-10	Occupational Information	4-8*	--	4-a, c	I-4	• Test contains some items with sex stereotyping; • Further research needed on internal consistency measures; • Needs to be reduced in length; • Forms overlap so that discrimination ability is lessened between grade levels.

---

[a] Table II, Part IV.

[b] Table III, Part IV.

[c] Table IV Part V

ASSESSMENT OF CAREER DEVELOPMENT

Authors: Dale Prediger, John Roth, Bert Westbrook

Publisher: Houghton Mifflin
One Beacon Street
Boston, Massachusetts 02107

Grade Levels: 8 - 11

Administration Time: 125 minutes (three sessions)

Administration Format: Written

Scoring: Cannot be scored locally except by special arrangement: summary data for five subscales and 42 additional items are reported.

Cost: Test Booklets: 35 - $13.95
Answer Sheets: 35 - $ 3.75

Remarks:

- Subscore key

Subscore 1 - Occupational characteristics
Subscore 2 - Occupational requirements
Subscore 3 - Exploratory occupational experiences, general
Separate scores for each of the following clusters:

- Social/Health and Personal Services
- Business Sales and Management
- Business Operations
- Technologies and Trades
- Natural, Social and Medical Sciences
- Creative and Applied Arts

Subscore 4 - Career Planning Knowledge
Subscore 5 - Career Planning Involvement

- Answer sheet provides for collecting responses for up to 19 locally constructed questions.

- Should not be used below 8th grade except for groups having above average verbal ability.

CAREER DEVELOPMENT INVENTORY

Authors:	Donald E. Super Martin J. Bohn, Jr. David J. Forrest	Jean Pierre Jordaan Richard H. Lindeman Albert S. Thompson
Publisher:	Available from: Donald E. Super, Teachers College Columbia University, New York, N. Y. 10027	
Grade Levels:	8 - 12	
Administration Time:	30 minutes for average student; all students should complete inventory in one class period.	
Administration Format:	Written.	
Scoring:	Can be scored locally or commercially (by advance arrangement to assure use of appropriate answer sheets) by the National Training and Evaluation Center, 135 East 65th St., New York, NY 10021.	
Costs:	$2.50 for specimen set; permission for local reproduction prior to publication included.	

Remarks:

- Scale key:

  Scale A: Planning Orientation
  Scale B: Resources for Exploration
  Scale C: Information and Decision Making

- There is sex stereotyping in a few items, but empirical studies show there is no sex bias; revision to remove stereotyping is in process.

CAREER MATURITY INVENTORY

Author: John O. Crites

Publisher: CTB/McGraw Hill
Del Monte Research Park
Monterey, California 93940

Grade Levels: 7 - 12; items should be presented orally to those students who are not able to read at the sixth grade level.

Administrative Time: 2-1/2 hours total; each of the six parts requires approximately 25 minutes.

Administrative Format: Written; however, for those not reading at 6th grade level, items may be presented orally.

Scoring: Can be scored locally or commercially.

Costs: Test booklets (Package of 35) - $20.00
Answer sheets (Package of 35) - $ 4.00

Remarks:

- Part key

    A - Attitude Scale (Attitude Toward Work)
    B - Competence Tests:
        Part 1 - Knowing Yourself (Self-Appraisal)
        Part 2 - Knowing About Jobs (Occupational Information)
        Part 3 - Choosing a Job (Goal Selection)
        Part 4 - Looking Ahead (Planning)
        *Part 5 - What Should They Do? (Problem Solving)

---

*Part 5 was not recommended by the review panel due to disagreement regarding the scoring of some of the items.

- Subscores on the following scales for the Attitude Scale, with new normative data, will be available in January 1976:

  Part 1: Involvement in Career Decision Making
  Part 2: Independence in Career Decision Making
  Part 3: Conceptions of Career Decision Making
  Part 4: Orientation to Career Decision Making
  Part 5 Factors in Career Decision Making

- Minor degree of sex stereotyping; no clear evidence of sex bias.

COGNITIVE QUESTIONNAIRE FOR CAREER EDUCATION

Authors: Billie T. Rader, Karin Nelson, Scott G. Anderson, John D. Skinkle

Publisher: Minnesota Research Coordinating Unit for Vocational Education
University of Minnesota
Minneapolis, Minnesota 55455

Grade Levels: 1 - 3; 4 - 6; 7 - 9

Administration Time: 30 - 45 minutes

Administration Format:
Grades 1 - 3: items read to student by teacher
Grades 4 - 6: student reads and responds to items
Grades 7 - 9: student reads and responds to items

Scoring: Hand scoring or optical scoring

Costs:

K - 3:	1 booklet which tests 10 students	$3.00
4 - 6:	1 booklet	.50
7 - 9:	1 booklet	.50
Manual		1.00
Specimen Sets		4.00

Remarks:

- Test measures areas of occupational information.
- Some items contain sex-ethnic group stereotyping.
- Some items have incorrect factual content.
- Overlapping forms cause poor discrimination between grade levels.

COOPERSMITH SELF-ESTEEM INVENTORY

Author:	Stanley Coopersmith	
Publisher:	Self Esteem Institute 1736 Stockton Street San Francisco, California 94133	
Grade Levels:	Age 9 to adult	
Administration Time:	Suggested time for Form A: 12 minutes Suggested times for Forms B & C: 6 minutes	
Administration Format:	The SEI has three forms: Forms A, B, and C. Form C (25 items) was designed for older (adult) groups. Form B is a short form revision of Form A (25 items vs. 58 items). Subjects read items and check response as either "like me" or "unlike me."	
Scoring:	Hand scoring	
Costs:	SEI Form A:	$30/100
	SEI Forms B & C:	$28/100
	Scoring Keys:	$ 1.00 each

Remarks:

- Test measures self esteem.
- The panel recommended that only the long form (Form A) be used.

DIFFERENTIAL APTITUDE TESTS with Career Planning Program

Authors:	George K. Bennett Harold G. Seashore Alexander G. Wesman Donald E. Super (Career Planning Program)
Publisher:	The Psychological Corporation 757 - 3rd Avenue New York, N. Y. 10017
Grade Levels:	8 - 12
Administration Time:	235 minutes
Administration Format:	Written
Scoring:	Can be scored locally or commercially.
Costs:	Test booklets (package of 25) - $18.50 Answer sheets (package of 50) - $12.50* Answer sheets (package of 50) - $11.00**

Remarks:

- Career Planning Program is currently formulated on an individual student basis; publisher is prepared to offer a group summary of findings.
- Might be utilized in lieu of achievement testing for outcomes 2-a, 2-b, Table III.

---

* IBM and Op Scan

** NCR

EMPLOYMENT READINESS SCALE

Author: Anthony M. Alfano

Publisher: Information may be obtained through author at:
Department of Counseling & Student Personnel Services
College of Education, Aderhold Bldg.
University of Georgia
Athens, Georgia 30602

Ages/Grade Levels: Adults/those seeking or having regular employment.
High School (11th and 12th grades)

Administration Time: Approximately 10-15 minutes

Administration Format: Self-administered

Scoring: Scored by hand

Cost: A copyright has been granted to the author. He has stated that fellow professionals will be granted the limited right to reproduce the scale. No price has been established thus far.

Remarks:

- This test measures one's readiness for employment by looking at one's work values.
- Test is especially applicable for high school students going to work upon graduation.
- Test needs further psychometric development.
- Low correlations among some test items.

HOW I SEE MYSELF SCALE

Author: Ira J. Gordon
2900 S.W. 2nd Court
Gainesville, Florida 32601

Grade Levels: 3 - 12

Administration Time: Untimed; 40 item elementary form and 42 item secondary form. No approximate time suggested. Probably 20-30 minutes.

Administration Format: Two forms; elementary and secondary. It is recommended that for younger students (grades 3, 4) each scale item be read separately and that each child understands the words and rating system. For older children, directions are given and each child reads and responds to each item.

Scoring: Hand scored; or
Can be scored by the:
Florida Educational Research and Development Council
College of Education
University of Florida
Gainesville, Florida 32601
$.25 per student

Cost: Tests: $ .05
Manual: $1.00

Remarks:

- Test measures self-esteem.
- One may encounter some scoring difficulties.
- Reliability of factor structures at the elementary level are questionable.

NEW MEXICO CAREER EDUCATION TEST SERIES

Authors: Charles C. Healy Stephen P. Klein

Publisher:
Monitor
P. O. Box 2337
Hollywood, California 90028

Grade Levels: 9 - 12

Administration Time: Tests can be administered separately or as a total battery.

	Minutes
Career Development Test	20
Career Oriented Activities Checklist	20
*Knowledge of Occupations Test	20
Job Application Procedures Test	20
Career Planning Test	20

Administration Format: Written

Scoring: Machine or hand scored.

Cost: For each subtest:

Test Booklet:	35 - $8.50; 100 - $22.50
Answer Sheet:	35 - $2.00; 100 - $ 4.50
Scoring Stencil	$1.00 each

Remarks:

- Test consists of five different subtests
    *Knowledge of Occupations Test
    Career Development Test was considered "promising."
    Career Oriented Activities Checklist was recommended.
    Job Application Procedures Test was considered "promising."
    Career Planning Test was considered "promising."

---

* Knowledge of Occupations Test was not recommended.

ORIENTATION TO CAREER CONCEPTS SERIES

Authors:	Barbara Fulton Robert Tolsma
Publisher:	Evaluative Research Associates, Inc. 8444 Florissant Road St. Louis, Missouri 63121
Grade Levels:	Designed primarily for grades 4 - 8. However, many of the tests can be used through grade 12.
Administration Time:	Three hours are required for the total battery of ten subtests. For any one subtest the time ranges from 10 - 30 minutes with most tests averaging about 20 minutes.
Administration Format:	The test series consists of a battery of ten 20-item tests. Each item has a five multiple choice alternative. Because the responses to some items include technical terms, students below grade 7 may have reading problems. For students below grade 7, the administrator may read the test aloud to the students.
Scoring:	Scoring can either be done by hand or machine services provided by Evaluative Research Associates.
Costs:	In packages of 35: $6.50 for each of the ten subtests. Any combination of two tests (pkgs of 35), $11.00. Total series (in pkg of 35), $36.00.

Remarks:

- Subtest Key

Test 1	Work Awareness	Test 6	Occupational Tools
Test 2	Worker Activities	Test 7	Work Stories
Test 3	Vocational Vocabulary	Test 8	Working Conditions
*Test 4	Absurdities	Test 9	Occupational Training
Test 5	Occupational Similarities	Test 10	Worker Earnings

- Test measures occupational information.
- Some items contain sex stereotyping.
- Needs to be reduced in length.
- Overlapping forms cause poor discrimination between grade levels.
- Further research needed on internal consistency measures.

* Test 4 not recommended.

THE PIERS-HARRIS CHILDREN'S SELF-CONCEPT SCALE

Authors:	Ellen V. Piers Dale B. Harris
Publisher:	Counselor Recordings and Tests Box 6184 Acklen Station Nashville, Tennessee 37212
Grade Levels:	3 - 12
Administration Time:	15 - 20 minutes
Administration Format:	This test consists of 80 first person declarative statements to which the student responds on an answer sheet, by circling a "yes" or "no."
Scoring:	Hand scored with scoring key.

Costs:

Tests:	$ .20 each
Scoring Key:	.50 each
Manual:	1.00 each

Remarks:

- This test measures self-esteem.
- Further data needs to be collected on subscores suggested by the factor analysis.

SELF-OBSERVATION SCALES

Authors:	Jack Stenner William Katzenmeyer
Publisher:	National Testing Service 2526 Erwin Road Durham, North Carolina 27705
Grade Levels:	K - 4, 5 - 12
Administration Time:	20 - 25 minues
Administration Format:	Written
Scoring:	Cannot be scored locally.
Costs:	Test booklets and answer sheets, and standard scoring in sets of 30.

1-5 sets	-	$35 per set
6-17 sets	-	32 per set
18-34 sets	-	21.60 per set
35-200 sets	-	17.40 per set

Remarks:

Subscale Key:

K-4 (Primary Level)

Subscale 1 - Self Acceptance
Subscale 2 - Social Maturity
Subscale 3 - School Affiliation
Subscale 4 - Self Security
*Subscale 5 - Achievement Motivation

---

* Subscales 5 (K-4) and 8 (5-12) were not recommended by the review panel due to insufficient psychometric data.

5-12 (Intermediate Level)

Subscale 1 - Self Acceptance
Subscale 2 - Self Security
Subscale 3 - Social Maturity
Subscale 4 - Social Confidence
Subscale 5 - School Affiliation
Subscale 6 - Teacher Affiliation
Subscale 7 - Peer Affiliation
*Subscale 8 - Achievement Motivation

- This test was recommended by the review panel only on the condition that the publisher provide users with information regarding the scoring keys and weights for the subscales.

- Local review for sensitive items was recommended by the review panel.

- The review panel recommended further work by the publisher on validation.

- There is a Spanish version of the SOS which was not reviewed by the panel.

YOUTH INVENTORY

Authors:	Hermann H. Remmers Benjamin Shimberg
Publisher:	Scholastic Testing Service, Inc. 480 Meyer Road Bensenville, Illinois 60106
Grade Levels:	7 - 12
Administration Time:	About 30 - 35 minutes
Administration Format:	Written. Student reads statement and checks box.
Scoring:	Hand or machine scored.
Costs:	Booklets in packages of 35:

1-15 pkgs.	-	$7.30 per pkg.
16-30 pkgs.	-	$6.60 per pkg.
over 31 pkgs.	-	$5.90 per pkg.

Answer sheets: $7.00 per pkg. of 50

General Manual: $2.00 per copy

Remarks:

- This test can be used to measure the areas of self-esteem, career awareness, and career attitudes.

IV. CAREER EDUCATION STUDENT OBJECTIVES

These sets of career education objectives were selected from among the many available because they are quite widely known in school districts across the country.

TABLE II

GENERAL CAREER EDUCATION LEARNER OUTCOMES: USOE POLICY STATEMENT*

Career Education seeks to produce individuals who, when they leave school (at any age or grade level), are:

1. Competent in basic academic skills required for adaptability in our rapidly changing society.

2. Equipped with good work habits.

3. Capable of choosing and who have chosen a personally meaningful set of work values that foster in them a desire to work.

4. Equipped with career decision-making skills, job hunting skills, and job getting skills.

5. Equipped with vocational personal skills at a level that will allow them to gain entry into and attain a degree of success in the occupational society.

6. Equipped with career decisions based on the widest possible set of data concerning themselves and their educational-vocational opportunities.

7. Aware of means available to them for continuing and recurrent education once they have left the formal system of schooling.

8. Successful in being placed in a paid occupation, in further education, or in a vocation consistent with their current career education.

9. Successful in incorporating work values into their total personal value structure in such a way that they are able to choose what, for them, is a desirable lifestyle.

---

*Kenneth B. Hoyt, An Introduction to Career Education. A Policy Paper of the U.S. Office of Education, 1975, pp. 10-11.

TABLE III

GENERAL STUDENT OUTCOME AREAS AND SPECIFIC OBJECTIVES APPLICABLE FOR ROUND II PROJECTS FUNDED UNDER PART D OF VOCATIONAL EDUCATION ACT*

General Outcome Areas	Specific Objectives
1. Self Awareness	a. Students will identify and describe their own current abilities and limitations.
	b. Students will identify and describe their own current interests and values.
	c. Students will endorse positive attitudes toward themselves.
2. Competency in Academic/Vocational Skills	a. Students will demonstrate generally useful numerical skills.
	b. Students will demonstrate generally useful communication skills.
	c. Students will demonstrate generally useful information processing skills.
	d. Students will demonstrate generally useful decision-making skills.
	e. Students will demonstrate generally useful interpersonal skills.
3. Set of Work Values	a. Students will identify the bases of various work values.
	b. Students will endorse positive attitudes toward paid and unpaid work.
4. Awareness of and Knowledge about Work	a. Students will identify the major duties and required abilities associated with different types of paid and unpaid work.
	b. Students will distinguish between differences in work conditions and life styles associated with different types of paid and unpaid work.
	c. Students will distinguish between entry requirements for major types of paid and unpaid work.
	d. Students will identify the impact of social and technological change on paid and unpaid work.
	e. Students will identify the important factors that affect work success and satisfaction
5. Career Decision Making Skills	a. Students will associate their own abilities and limitations with possible success in present or future paid and unpaid work.
	b. Students will relate their personal interests and values to types of paid and unpaid work and their associated life styles.
	c. Students will identify, locate, and utilize sources of information to solve career decision-making problems.
	d. Students will determine the potential for future advancement/personal growth in work of their choosing.
	e. Students will identify the sequence of steps to be taken and the factors to be considered in career planning.
	f. Students will demonstrate active involvement in career decision making.
6. Work Habits	a. Students will plan work effectively.
	b. Students will adapt to varied work conditions.
	c. Students will endorse a positive attitude towards the concept of quality in relation to a work task.
	d. Students will endorse a positive attitude towards conservation.
	e. Students will endorse a positive attitude towards responsibility for their own behavior and accomplishment of self-imposed tasks.
	f. Students will demonstrate a desire for continuous learning, both in school and out.

TABLE III-- cont.

General Outcome Areas		Specific Objectives
7. Work-seeking and Work-getting skills	a.	Students will identify, locate, and utilize sources that contain information about existing paid and unpaid possibilities.
	b.	Students will demonstrate skills required in applying for and accepting work.
8. Successful Placement of Students Upon Leaving Educational System	a.	Students will be placed in a paid occupation, in further education, or in unpaid work that is consistent with their current career plans.
9. Awareness of Means Available for Continued Education	a.	Students will identify sources of additional education in major types of paid and unpaid work.
	b.	Students will identify means to support additional education for themselves in major types of paid and unpaid work.

* Developed in June 1974 by US Office of Education and included in draft version of this guide, August 15, 1974.

TABLE IV

STUDENT OUTCOME OBJECTIVES OF THE EXPERIENCE BASED CAREER EDUCATION PROGRAMS (EBCE)*

I. CAREER DEVELOPMENT SKILLS AND KNOWLEDGE

I-1. The student can integrate (relate) information about occupations with information about self.

I-2. The student can locate and use information about occupations.

I-3. The student has a positive attitude (orientation) toward career planning.

I-4. The student knows the functions, characteristics, and requirements of a broad range of self-selected occupations.

I-5. The student knows some of the factors associated with selected occupations that contribute to job success and job satisfaction.

I-6. The student will demonstrate that he/she has made an informed decision regarding his or her post high school educational/vocational plans.

I-7. The student can obtain employment information, complete job applications, take interviews, write letters of application, prepare a resume, etc.

I-8. Students who have tentatively selected a career area can begin to acquire some of the related job entry skills and experience.

II. SELF KNOWLEDGE: INTERESTS, ABILITIES, AND VALUES

II-1. The student can accurately demonstrate awareness and understanding of his or her own current interests, abilities, values, and limitations relevant to career goal selection and achievement, and recognize that these may change with further education or experience.

---

* These objectives were used in the design of the evaluation of EBCE by the Educational Testing Service for the National Institute of Education, winter 1975.

III. READING SKILLS

III-1. The student can read selections from a newspaper or other popular periodical and (1) recognize the main point(s), (2) recognize the author's purpose, and (3) locate specific facts and details.

III-2. The student can read and comprehend materials pertaining to his or her areas of career involvement, such as instructions, manuals, forms, parts lists, and technical articles.

III-3. The student can read and comprehend materials appropriate to his or her vocational and recreational interests.

III-4. The student can read selections required for educational or occupational advancement and (1) define the author's purpose and support that definition with evidence, (2) identify and explain different levels of meaning included in the selection, (3) identify biases with supporting evidence, (4) extend interpretation beyond the printed information, and (5) recognize and describe different writing styles.

IV. PROBLEM SOLVING SKILLS

IV-1. The student can define his/her problem by identifying a need or a discrepancy between where he or she is and where he or she wants to be. This can be in a personal, group, societal, academic, and/or career situation.

IV-2. The student can use a variety of sources and techniques of data gathering.

IV-3. The student can propose or generate alternative solutions, anticipate consequences of various actions, and implement a course of action.

V. ORAL COMMUNICATION

V-1. The student will demonstrate an ability to communicate orally both ideas and feelings in a manner that is effective and appropriate to various situations (social, school, or work).

V-2. The student will demonstrate an ability to listen effectively.

VI. WRITING SKILLS

VI-1. The student can express in writing ideas and feelings so that most people can understand what was stated.

VI-2. The student can write clearly and correctly the materials pertaining to his or her areas of career involvement (e.g., reports, orders, records, and forms).

VI-3. The student can write letters, descriptions, and reports required in normal daily living.

VII. INTERPERSONAL SKILLS

VII-1. The student can effectively participate in peer and adult interactions based on appropriate role relationships and obligations, acceptance of the validity of individual rights and perceptions, and ability to contribute to the resolutions of conflicts resulting from differing personal needs and values.

VII-2. The student will demonstrate the ability to cooperate with others as a means of attaining goals.

VIII. BASIC QUANTITATIVE SKILLS

VIII-1. The student will demonstrate an ability to comprehend and interpret information presented numerically and graphically as found in such media as newspapers and weekly news magazines.

VIII-2. The student will demonstrate correct performance of arithmetic operations necessary for successful daily living, such as (1) making and receiving change, (2) modifying recipe quantities, (3) measuring items, (4) doing comparison shopping, and (5) generally dealing with weights, measures, calendars, clocks, etc.

VIII-3. The student will demonstrate correct performance of mathematical operations necessary for his/her chosen career, or for meeting the requirements for continued study if a continuation of formal education is chosen as the next stage in career development.

IX. MATURATION SKILLS

IX-1. The student will demonstrate the ability to use direct sources (i.e., observations or interviews with relevant people) in greater proportion to indirect sources (i.e., books written about a topic) in gathering information for reports and projects.

IX-2. The student will demonstrate the ability to conduct conversations with an adult that reveals the student's self-confidence, ability to discuss a fixed topic for a reasonable amount of time, and an understanding of the other person's message and feelings.

IX-3. The student will demonstrate the ability to cooperate with adults and assume responsibility for carrying out tasks which he/she agrees to complete.

IX-4. The student will demonstrate an increase in behaviors that reveal a tolerance for people who are different in ideas or background than himself/herself, an openness to change and a willingness to trust others when circumstances warrant.

V. LIST OF PANEL MEMBERS

Dr. Nancy Burton

Dr. Burton is currently Scoring Coordinator, Department of Research and Analysis, National Assessment of Educational Progress, Education Commission of the States. Formerly, she was responsible for development of the National Assessment's effort in Career and Occupational Development.

Dr. William Coffman

Dr. Coffman is currently Lindquist Professor of Education and Director of the Iowa Testing Programs at the University of Iowa. He has been a member of the Analysis Advisory Committee for the National Assessment of Educational Progress, Past President of the National Council on Measurement in Education, a fellow of the APA, and a member of AERA. He has served as a research assistant for the Horace-Mann Lincoln Institute at Teachers College, Columbia University; Associate Professor of Psychology at the Oklahoma State University; and Director of various divisions of the Educational Testing Service. Dr. Coffman is the author of numerous articles and publications.

Dr. John O. Crites

Dr. Crites is currently Professor of Psychology at the University of Maryland. He is a member of the APA (President, division 17); AERA; and American Personnel and Guidance Association. He has served as a counseling psychologist at the University of Texas, Professor of Psychology at the University of Iowa, and Director of the University of Iowa Counseling Center. He is the author of the Career Maturity Inventory and numerous publications.

Dr. Lois-ellin Datta

Dr. Datta is currently Deputy Associate Director of Career Education for National Institute of Education. She has served as a Senior Staff Fellow at NIH; a research psychologist for the National Institute of Mental Health; chief evaluator for project Head Start; and Chief, Childhood Research and Development, Office of Child Development.

Dr. Edward Lareau

Dr. Lareau is currently an Associate Director for Research, Research Coordinating Unit, Admiral Peary Vocational-Technical School, Ebensburg, Pennsylvania. He is also Director of the Career Education Project sponsored by the area vocational-technical school.

Dr. Dale Prediger

Dr. Prediger is currently an Assistant Director, Research Services Division and Assistant Director, Developmental Research Department, of the Americal College Testing Program. He directed the development of the Assessment of Career Development.

Dr. Donald Super

Dr. Super is currently Professor of Psychology at Teachers College, Columbia University. He is a member of the APA (past President - - Division of Counseling Psychology); past President of the American Personnel and Guidance Association; past President of National Vocational Guidance Association, and Vice President of International Association for Education and Vocational Guidance. He is the author of numerous tests (e.g., Work Values Inventory), articles, and publications.

Dr. Bert Westbrook

Dr. Westbrook is currently an Associate Professor of Psychology at North Carolina State University. He is also a consultant for the National Assessment of Educational Progress; the Appalachian Educational Laboratory; the North Carolina Department of Youth Development, and the American College Testing Program. He is a member of the APA; AERA, National Council for Measurement in Education; and the American Personnel and Guidance Association. He has taught in the public school system, and is the author of numerous articles and publications.